

BUNDESREPUBLIK DEUTSCHLAND



Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 199 42 123.4

Anmeldetag: 03. September 1999

Anmelder/Inhaber: BASF Aktiengesellschaft, Ludwigshafen/DE

Bezeichnung: Codierende Gene neuer Proteine von Corynebacterium Glutamicum

IPC: C 07 K, C 12 N

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 2. Juli 2002
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag



Faust

CERTIFIED COPY OF
PRIORITY DOCUMENT

CODIERENDE GENE NEUER PROTEINE VON *CORYNEBACTERIUM GLUTAMICUM*

Hintergrund der Erfindung

5

Bestimmte Produkte und Nebenprodukte von natürlich vorkommenden Stoffwechselprozessen in Zellen werden in vielen Industriezweigen verwendet, einschließlich der Nahrungsmittel-, Futtermittel-, Kosmetik- und pharmazeutischen Industrie. Diese Moleküle, die ge-

10 meinsam als "Feinchemikalien" bezeichnet werden, umfassen organische Säuren, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Nukleotide und Nukleoside, Lipide und Fettsäuren, Dirole, Kohlehydrate, aromatische Verbindungen, Vitamine und Co-

15 faktoren sowie Enzyme. Ihre Produktion erfolgt am besten mittels Anzucht von Bakterien im Großmaßstab, die entwickelt wurden, um große Mengen des jeweils gewünschten Moleküls zu produzieren und sezernieren. Ein für diesen Zweck besonders geeigneter Organismus ist *Corynebacterium glutamicum*, ein gram-positives, nicht-pathogenes Bakterium. Über Stammselektion ist eine Reihe von Mutanten-

20 stämmen entwickelt worden, die ein Sortiment wünschenswerter Verbindungen produzieren. Die Auswahl von Stämmen, die hinsichtlich der Produktion eines bestimmten Moleküls verbessert sind, ist jedoch ein zeitaufwendiges und schwieriges Verfahren.

25 Zusammenfassung der Erfindung

Diese Erfindung stellt neuartige Nukleinsäuremoleküle bereit, die sich zur Identifizierung oder Klassifizierung von *Corynebacterium glutamicum* oder verwandten Bakterienarten verwenden lassen. C.

30 *glutamicum* ist ein gram-positives, aerobes Bakterium, das in der Industrie für die Produktion im Großmaßstab einer Reihe von Feinchemikalien und auch zum Abbau von Kohlenwasserstoffen (bspw. beim Überlaufen von Rohöl) und zur Oxidation von Terpenoiden gemeinhin verwendet wird. Die Nukleinsäuremoleküle können daher zum

35 Identifizieren von Mikroorganismen eingesetzt werden, die sich zur Produktion von Feinchemikalien, bspw. durch Fermentationsverfahren, verwenden lassen. *C. glutamicum* selbst ist zwar nicht-pathogen, jedoch ist es mit anderen *Corynebacterium*-Arten, wie *Corynebacterium diphtheriae* (dem Erreger der Diphtherie) verwandt,

40 die bedeutende Pathogene beim Menschen sind. Die Fähigkeit, das Vorhandensein von *Corynebacterium*-Arten zu identifizieren, kann daher auch eine signifikante klinische Bedeutung haben, z.B. bei diagnostischen Anwendungen. Diese Nukleinsäuremoleküle können zudem als Bezugspunkte zur Kartierung des *C. glutamicum*-Genoms oder

45 von Genomen verwandter Organismen dienen.

Diese neuen Nukleinsäuremoleküle codieren Proteine, die hier als Marker- und Feinchemikalienproduktions- (MCP-) Proteine bezeichnet werden. Diese MCP-Proteine können bspw. direkt oder indirekt an der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* beteiligt sein. Die erfindungsgemäßen MCP-Proteine können auch am Abbau von Kohlenwasserstoffen oder an der Oxidation von Terpenoiden beteiligt sein. Diese Proteine lassen sich zur Identifikation von *Corynebacterium glutamicum* oder von zu *C. glutamicum* verwandten Organismen verwenden; das Vorliegen eines für *C. glutamicum* und verwandte Arten spezifischen MCP-Proteins in einem Proteingemisch kann auf das Vorliegen eines dieser Bakterien in der Probe hindeuten. Ferner können diese MCP-Protein Homologa in Pflanzen oder Tieren aufweisen, die an einem Erkrankungszustand oder einem Leiden beteiligt sind; so können diese Proteine als nützliche pharmazeutische Ziele für das Arzneimittel-Screening und die Entwicklung therapeutischer Verbindungen dienen.

Aufgrund der Verfügbarkeit von in *Corynebacterium glutamicum* verwendbaren Klonierungsvektoren, wie bspw. offenbart in Sinskey et al., US-Patent Nr. 4 649 119, und von Techniken zur genetischen Manipulation von *C. glutamicum* und den verwandten *Brevibacterium*-Arten (z.B. *lactofermentum*) (Yoshihama et al., J. Bacteriol. 162:591-597 (1985); Katsumata et al., J. Bacteriol. 159:306-311 (1984); und Santamaria et al. J. Gen. Microbiol. 130:2237-2246 (1984)), lassen sich die erfindungsgemäßen Nukleinsäuremoleküle zur genetischen Manipulation dieses Organismus verwenden, um die Produktion einer oder mehrerer Feinchemikalien zu modulieren. Diese Modulation kann aufgrund einer direkten Auswirkung der Manipulation eines erfindungsgemäßen Gens oder aufgrund einer indirekten Auswirkung einer solchen Manipulation erfolgen. Bspw. kann man durch Modifikation der Aktivität eines Proteins, das an der Biosynthese oder am Abbau einer Feinchemikalie beteiligt ist, (d.h. durch Mutagenese des entsprechenden Gens) die Fähigkeit der Zelle zur Synthese oder zum Abbau dieser Verbindung direkt modulieren, wodurch die Ausbeute und/oder Effizienz der Produktion der Feinchemikalie moduliert wird. Ebenso kann man durch Modulation der Aktivität eines Proteins, das einen Feinchemikalien-Stoffwechselweg reguliert, direkt beeinflussen, ob die Produktion der gewünschten Verbindung hoch- oder herunterreguliert wird, was beides die Ausbeute oder Effizienz der Produktion der Feinchemikalie von der Zelle moduliert.

Die indirekte Modulation der Feinchemikalienproduktion kann auch durch Modifikation der Aktivität eines erfindungsgemäßen Proteins (d.h. durch Mutagenese des entsprechenden Gens) erfolgen, so daß die Fähigkeit der Zelle, zu wachsen und sich zu teilen oder lebensfähig und produktiv zu bleiben, insgesamt erhöht ist. Die

Produktion von Feinchemikalien aus *C. glutamicum* wird gewöhnlich durch Fermentationskultur im Großmaßstab dieser Mikroorganismen erzielt, Bedingungen, die für das Wachstum und die Zellteilung häufig suboptimal sind. Durch Verändern eines erfindungsgemäßen

5 Proteins (z.B. eines Streßreaktionsproteins, eines Zellwandproteins oder eines Proteins, das am Stoffwechsel von Verbindungen beteiligt ist, die für das Auftreten von Zellteilung und -Wachstum nötig sind, wie Nukleotide und Aminosäuren), so daß ein besseres Überleben, Wachsen und Vermehren in diesen Bedingungen mög-

10 lich ist, kann es möglich sein, die Anzahl und die Produktivität dieser veränderten *C. glutamicum*-Zellen in Kulturen im Großmaßstab zu steigern, was wiederum zu gesteigerten Ausbeuten und/oder zu gesteigerter Effizienz der Produktion einer oder mehrerer gewünschter Feinchemikalien führen sollte. Ferner sind die Stoff-

15 wechselwege einer Zelle notwendigerweise voneinander abhängig und co-reguliert. Durch Ändern der Aktivität irgendeines Stoffwechselwegs in *C. glutamicum* (d.h. durch Ändern der Aktivität eines der erfindungsgemäßen Proteine, das an einem solchen Weg beteiligt ist) ist es möglich, gleichzeitig die Aktivität oder Regula-

20 tion eines anderen Stoffwechselwegs in diesem Mikroorganismus zu ändern, der direkt an der Synthese oder am Abbau einer Feinchemikalie beteiligt sein kann.

Diese Erfindung stellt neue Nukleinsäuremoleküle bereit, die Pro-

25 teine codieren, die hier als MCP-Proteine bezeichnet werden und bspw die Produktion oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen dienen können. Nukleinsäuremoleküle, die ein MCP-Protein codie-

30 ren, werden hier als MCP-Nukleinsäuremoleküle bezeichnet. Bei einer bevorzugten Ausführungsform kann das MCP-Protein die Produktion oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen dienen. Beispiele für

35 solche Proteine sind diejenigen, die von den in Tabelle 1 angegebenen Genen codiert werden.

Ein Aspekt der Erfindung betrifft folglich isolierte Nukleinsäuremoleküle (bspw. cDNAs), umfassend eine Nukleotidsequenz, die

40 ein MCP-Protein oder biologisch aktive Abschnitte davon codiert, sowie Nukleinsäurefragmente, die sich als Primer oder Hybridisierungssonden zum Nachweis oder zur Amplifikation von MCP-codierenden Nukleinsäure (bspw. DNA oder mRNA) eignen. Bei besonders bevorzugten Ausführungsformen umfaßt das isolierte Nukleinsäuremo-

45 lekül eine der in Anhang A aufgeführten Nukleotidsequenzen oder den codierenden Bereich einer dieser Nukleotidsequenzen oder ein Komplement davon. In anderen besonders bevorzugten Ausführungs-

formen umfaßt das erfindungsgemäße isolierte Nukleinsäuremolekül eine Nukleotidsequenz, die an eine in Anhang A angegebene Nukleotidsequenz hybridisiert oder mindestens zu etwa 50%, vorzugsweise mindestens zu etwa 60%, stärker bevorzugt mindestens etwa 70%, 5 80% oder 90% und noch stärker bevorzugt mindestens etwa 95%, 96%, 97%, 98%, 99% oder noch homologer dazu ist, oder einen Abschnitt davon. In anderen bevorzugten Ausführungsformen codiert das isolierte Nukleinsäuremolekül eine der in Anhang B aufgeführten Aminosäuresequenzen. Die bevorzugten erfindungsgemäßen MCP-Proteine 10 besitzen ebenfalls vorzugsweise mindestens eine der hier beschriebenen MCP-Aktivitäten.

Bei einer weiteren Ausführungsform codiert das isolierte Nukleinsäuremoleküle ein Protein oder einen Abschnitt davon, wobei das 15 Protein oder sein Abschnitt eine Aminosäuresequenz enthält, die zu einer Aminosäuresequenz in Anhang B hinreichend homolog ist, bspw. zu einer Aminosäuresequenz in Anhang B derart hinreichend homolog ist, daß das Protein oder sein Abschnitt eine MCP-Aktivität behält. Vorzugsweise behält das von dem Nukleinsäuremolekül 20 codierte Protein oder der Abschnitt davon die Fähigkeit, die Produktion oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* zu modulieren oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen zu dienen. Bei einer Ausführungsform ist das von dem Nukleinsäuremolekül co- 25 dierte Protein mindestens etwa 50%, vorzugsweise mindestens etwa 60% und stärker bevorzugt mindestens etwa 70%, 80% oder 90% und am stärksten bevorzugt mindestens etwa 95%, 96%, 97%, 98% oder 99% oder noch homologer zu einer Aminosäuresequenz in Anhang B (bspw. einer vollständigen Aminosäuresequenz, ausgewählt aus den 30 in Anhang B genannten Sequenzen). Bei einer weiteren bevorzugten Ausführungsform ist das Protein ein *C. glutamicum*-Vollängenprotein, das zu einer vollständigen Aminosäuresequenz in Anhang B (die von einem in Anhang A gezeigten offenen Leseraster codiert wird) im wesentlichen homolog ist.

35 Bei einer weiteren bevorzugten Ausführungsform stammt das isolierte Nukleinsäuremolekül aus *C. glutamicum* und codiert ein Protein (z.B. ein MCP-Fusionsprotein), das eine biologisch aktive Domäne umfaßt, die zu einer der Aminosäuresequenzen in Anhang B 40 zumindest zu etwa 50% oder stärker homolog ist und die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren, Kohlenwasserstoffe abbauen, Terpenoide oxidieren, als Ziel für Arzneimittelentwicklung dienen oder als Identifikationsmarker für *C. glutamicum* oder 45 verwandte Organismen dienen kann und auch heterologe Nukleinsäu-

resequenzen enthält, die ein heterologes Polypeptid oder regulatorische Bereiche codieren.

Bei einer weiteren Ausführungsform ist das isolierte Nukleinsäuremolekül mindestens 15 Nukleotide lang und hybridisiert unter stringenten Bedingungen an ein Nukleinsäuremolekül, das eine Nukleotidsequenz aus Anhang A umfaßt. Das isolierte Nukleinsäuremolekül entspricht vorzugsweise einem natürlich vorkommenden Nukleinsäuremolekül. Die isolierte Nukleinsäure codiert stärker bevorzugt ein natürlich vorkommendes *C. glutamicum*-MCP-Protein oder einen biologisch aktiven Abschnitt davon.

Ein weiterer Aspekt der Erfindung betrifft Vektoren, bspw. rekombinante Expressionsvektoren, die die erfindungsgemäßen Nukleinsäuremoleküle enthalten, und Wirtszellen, in die diese Vektoren eingebracht worden sind. Bei einer Ausführungsform wird diese Wirtszelle zur Herstellung eines MCP-Proteins verwendet, indem die Wirtszelle in einem geeigneten Medium gezüchtet wird. Das MCP-Protein kann dann aus dem Medium oder der Wirtszelle isoliert werden.

Ein weiterer Aspekt der Erfindung betrifft einen genetisch veränderten Mikroorganismus, bei dem ein MCP-Gen eingebracht oder verändert worden ist. Das Genom des Mikroorganismus ist bei einer Ausführungsform durch Einbringen eines erfindungsgemäßen Nukleinsäuremoleküls verändert worden, das die MCP-Wildtyp- oder die mutierte MCP-Sequenz als Transgen codiert. Bei einer anderen Ausführungsform ist ein endogenes MCP-Gen im Genom des Mikroorganismus durch homologe Rekombination mit einem veränderten MCP-Gen verändert, z.B. funktionell disruptiert, worden. Der Mikroorganismus gehört bei einer bevorzugten Ausführungsform zur Gattung *Corynebacterium* oder *Brevibacterium*, wobei *Corynebacterium glutamicum* besonders bevorzugt ist. Der Mikroorganismus wird in einer bevorzugten Ausführungsform auch zur Herstellung einer gewünschten Verbindung, wie einer Aminosäure verwendet, wobei Lysin besonders bevorzugt ist.

Ein weiterer Aspekt der Erfindung betrifft ein isoliertes MCP-Protein oder einen Abschnitt, bspw. einen biologisch aktiven Abschnitt, davon. Das isolierte MCP-Protein oder sein Abschnitt kann in einer bevorzugten Ausführungsform die Produktion oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen dienen. Bei einer weiteren bevorzugten Ausführungsform ist das isolierte MCP-Protein oder ein Abschnitt davon hinreichend homolog zu einer Aminosäuresequenz von Anhang B, so daß das Protein oder sein Abschnitt die Fähig-

keit behält, bspw. die Produktion oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* zu modulieren oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen zu dienen.

5

- Die Erfindung betrifft zudem ein isoliertes MCP-Proteinpräparat. Das MCP-Protein umfaßt bei bevorzugten Ausführungsformen eine Aminosäuresequenz aus Anhang B. Bei einer weiteren bevorzugten Ausführungsform betrifft die Erfindung ein isoliertes Vollängenprotein, das zu einer vollständigen Aminosäuresequenz aus Anhang B (welche von einem offenen Leseraster in Anhang A codiert wird) im wesentlichen homolog ist. Bei einer weiteren Ausführungsform ist das Protein mindestens zu etwa 50%, vorzugsweise mindestens zu etwa 60%, stärker bevorzugt mindestens etwa 70%, 80% oder 90% und am stärksten bevorzugt mindestens etwa 95%, 96%, 97%, 98%, oder 99% oder noch homologer zu einer vollständigen Aminosäuresequenz aus Anhang B. Das isolierte MCP-Protein umfaßt bei anderen Ausführungsformen eine Aminosäuresequenz, die zu mindestens etwa 50% oder stärker zu einer der Aminosäuresequenzen aus Anhang B homolog ist und die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren, Kohlenwasserstoffe abbauen, Terpenoide oxidieren, als Ziel für Arzneimittelentwicklung dienen oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen dienen kann.

25

- Das isolierte MCP-Protein kann alternativ eine Aminosäuresequenz umfassen, die von einer Nukleotidsequenz codiert wird, welche mit einer Nukleotidsequenz aus Anhang B, bspw. unter stringenten Bedingungen, hybridisiert oder mindestens zu etwa 50%, vorzugsweise mindestens zu etwa 60%, stärker bevorzugt mindestens etwa 70%, 80% oder 90% und noch stärker bevorzugt mindestens etwa 95%, 96%, 97%, 98% oder 99% oder noch homologer dazu ist. Bevorzugte Formen der MCP-Proteine haben ebenfalls vorzugsweise eine oder mehrere der hier beschriebenen biologischen Aktivitäten von MCP.

35

- Das MCP-Polypeptid oder ein biologisch aktiver Abschnitt davon kann mit einem Nicht-MCP-Polypeptid funktionsfähig verbunden werden, damit ein Fusionsprotein entsteht. Dieses Fusionsprotein hat bei bevorzugten Ausführungsformen eine andere Aktivität als das MCP-Protein allein. Bei anderen bevorzugten Ausführungsformen kann dieses Fusionsprotein die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen dienen. Die Integration dieses Fusionsproteins in eine Wirtszelle moduliert bei besonders bevor-

45

zugten Ausführungsformen die Produktion einer gewünschten Verbindung von der Zelle.

Ein weiterer Aspekt der Erfindung betrifft ein Verfahren zur Herstellung einer Feinchemikalie. Das Verfahren sieht die Anzucht einer Zelle vor, die einen Vektor enthält, der die Expression eines erfindungsgemäßen MCP-Nukleinsäuremoleküls bewirkt, so daß eine Feinchemikalie produziert wird. Dieses Verfahren umfaßt bei einer bevorzugten Ausführungsform zudem den Schritt der Gewinnung einer Zelle, die einen solchen Vektor enthält, wobei die Zelle mit einem Vektor transfiziert ist, der die Expression einer MCP-Nukleinsäure bewirkt. Dieses Verfahren umfaßt bei einer weiteren bevorzugten Ausführungsform zudem den Schritt, bei dem die Feinchemikalie aus der Kultur gewonnen wird. Die Zelle gehört bei einer besonders bevorzugten Ausführungsform zur Gattung *Corynebacterium* oder *Brevibacterium* oder ist aus den in Tabelle 3 angegebenen Stämmen ausgewählt.

Ein weiterer Aspekt der Erfindung betrifft Verfahren zur Modulation der Produktion eines Moleküls von einem Mikroorganismus. Diese Verfahren umfassen das Zusammenbringen der Zelle mit einer Substanz, die die MCP-Proteinaktivität oder die MCP-Nukleinsäure-Expression moduliert, so daß eine zellassoziierte Aktivität verglichen mit der gleichen Aktivität bei Fehlen der Substanz verändert wird. Die Zelle wird bei einer bevorzugten Ausführungsform hinsichtlich einer oder mehrerer *C. glutamicum*-MCP-Proteinaktivitäten moduliert, so daß die Ausbeute, Produktion und/oder Effizienz der Produktion einer gewünschten Feinchemikalie durch diesen Mikroorganismus verbessert wird. Die Substanz, die die MCP-Proteinaktivität moduliert, kann eine Substanz sein, die die MCP-Proteinaktivität oder die MCP-Nukleinsäure-Expression stimuliert. Beispiele für Substanzen, die die MCP-Proteinaktivität oder MCP-Nukleinsäureexpression stimulieren, umfassen kleine Moleküle, aktive MCP-Proteine und Nukleinsäuren, die MCP-Proteine codieren und in die Zelle eingebracht worden sind. Beispiele für Substanzen, die die MCP-Aktivität oder -Expression hemmen, umfassen kleine Moleküle und MCP-Antisense-Nukleinsäuremoleküle.

Ein weiterer Aspekt der Erfindung betrifft Verfahren zur Modulation der Ausbeuten, der Produktion und/oder der Effizienz der Produktion einer gewünschten Verbindung aus einer Zelle, umfassend das Einbringen eines MCP-Wildtyp- oder -Mutantengens in eine Zelle, das entweder auf einem gesonderten Plasmid bleibt oder in das Genom der Wirtszelle integriert wird. Die Integration in das Genom kann zufallsgemäß oder durch homologe Rekombination erfolgen, so daß das native Gen durch die integrierte Kopie ersetzt wird, was die Produktion der gewünschten Verbindung von der zu

modulierenden Zelle hervorruft. Diese Ausbeuten sind bei einer bevorzugten Ausführungsform erhöht. Bei einer weiteren bevorzugten Ausführungsform ist die Chemikalie eine Feinchemikalie, die in einer besonders bevorzugten Ausführungsform eine Aminosäure ist. Diese Aminosäure ist in einer besonders bevorzugten Ausführungsform L-Lysin.

Eingehende Beschreibung der Erfindung

- 10 Die vorliegende Erfindung stellt MCP-Nukleinsäure- und -Proteinmoleküle bereit, die zur Identifikation von *Corynebacterium glutamicum* oder verwandter Organismen, zur Kartierung des *C. glutamicum*-Genoms (oder des Genoms eines nah verwandten Organismus) oder zur Identifikation von Mikroorganismen verwendet werden können, die zur Produktion von Feinchemikalien, z.B. durch Fermentationsverfahren, verwendet werden können. Die von diesen Nukleinsäuren codierten Proteine können zur direkten oder indirekten Modulation der Produktion oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum*, als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen, zur Oxidation von Terpenoiden oder zum Abbau von Kohlenwasserstoffen oder als Ziele zur Entwicklung therapeutischer pharmazeutischer Verbindungen verwendet werden. Die erfindungsgemäßen Aspekte sind nachstehend weiter erläutert.

25

I. Feinchemikalien

- Der Begriff "Feinchemikalie" ist im Fachgebiet bekannt und beinhaltet Moleküle, die von einem Organismus produziert werden und in verschiedenen Industriezweigen Anwendungen finden, wie bspw., jedoch nicht beschränkt auf die pharmazeutische Industrie, die Landwirtschafts- und Kosmetikindustrie. Diese Verbindungen umfassen organische Säuren, wie Weinsäure, Itaconsäure und Diaminopimelinsäure, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Purin- und Pyrimidinbasen, Nukleoside und Nukleotide (wie bspw. beschrieben in Kunitake, A. (1996) Nucleotides and related compounds, S. 561-612, in Biotechnology Bd. 6, Rehm et al., Hrsg. VCH: Weinheim und den darin enthaltenen Zitaten), Lipide, gesättigte und ungesättigte Fettsäuren (bspw. Arachidonsäure), Diole (bspw. Propandiol und Butandiol), Kohlehydrate (bspw. Hyaluronsäure und Trehalose), aromatische Verbindungen (bspw. aromatische Amine, Vanillin und Indigo), Vitamine und Cofaktoren (wie beschrieben in Ullmann's Encyclopedia of Industrial Chemistry, Bd. A27, "Vitamins", S. 443-613 (1996) VCH: Weinheim und den darin enthaltenen Zitaten; und Ong, A.S., Niki, E. und Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations

- in Malaysia and the Society for Free Radical Research - Asien, abgehalten am 1.-3. Sept. 1994 in Penang, Malaysia, AOCS Press (1995)), Enzyme und sämtliche anderen von Gutcho (1983) in *Chemicals by Fermentation*, Noyes Data Corporation, ISBN: 0818805086
- 5 und den darin angegebenen Literaturstellen beschriebenen Chemikalien. Der Metabolismus und die Verwendungen bestimmter Feinchemikalien sind nachstehend weiter erläutert.

A. Metabolismus und Verwendungen von Aminosäuren

10

- Die Aminosäuren umfassen die grundlegenden Struktureinheiten sämtlicher Proteine und sind somit für die normalen Zellfunktionen in allen Organismen essentiell. Der Begriff "Aminosäure" ist im Fachgebiet bekannt. Die proteinogenen Aminosäuren, von denen
- 15 es 20 Arten gibt, dienen als Struktureinheiten für Proteine, in denen sie über Peptidbindungen miteinander verknüpft sind, wohingegen die nicht-proteinogenen Aminosäuren (von denen Hunderte bekannt sind) gewöhnlich nicht in Proteinen vorkommen (siehe Ulmann's Encyclopedia of Industrial Chemistry, Bd. A2, S. 57-97
- 20 VCH: Weinheim (1985)). Die Aminosäuren können in der optischen D- oder L-Konfiguration vorliegen, obwohl L-Aminosäuren gewöhnlich der einzige Typ sind, den man in natürlich vorkommenden Proteinen vorfindet. Biosynthese- und Abbauwege von jeder der 20 proteinogenen Aminosäuren sind sowohl bei prokaryotischen als auch euka-
- 25 ryotischen Zellen gut charakterisiert (siehe bspw. Stryer, L., Biochemistry, 3. Auflage, S. 578-590 (1988)). Die "essentiellen" Aminosäuren (Histidin, Isoleucin, Leucin, Lysin, Methionin, Phenylalanin, Threonin, Tryptophan und Valin), so bezeichnet, weil sie aufgrund der Komplexität ihrer Biosynthesen gewöhnlich mit
- 30 der Ernährung aufgenommen werden müssen, werden durch einfache Biosynthesewege in die übrigen 11 "nichtessentiellen" Aminosäuren (Alanin, Arginin, Asparagin, Aspartat, Cystein, Glutamat, Glutamin, Glycin, Prolin, Serin und Tyrosin) umgewandelt. Höhere Tiere besitzen die Fähigkeit, einige dieser Aminosäuren zu syn-
- 35 thetisieren, jedoch müssen die essentiellen Aminosäuren mit der Nahrung aufgenommen werden, damit eine normale Proteinsynthese stattfindet.

- Abgesehen von ihrer Funktion bei der Proteinbiosynthese sind
- 40 diese Aminosäuren interessante Chemikalien an sich, und man hat entdeckt, daß viele bei verschiedenen Anwendungen in der Nahrungsmittel-, Futter-, Chemie-, Kosmetik-, Landwirtschafts- und pharmazeutischen Industrie zum Einsatz kommen. Lysin ist nicht nur für die Ernährung des Menschen eine wichtige Aminosäure, son-
- 45 dern auch für monogastrische Tiere, wie Geflügel und Schweine. Glutamat wird am häufigsten als Geschmacksadditiv (Mononatriumglutamat, MSG) sowie weithin in der Nahrungsmittelindustrie ver-

- wendet, wie auch Aspartat, Phenylalanin, Glycin und Cystein. Glycin, L-Methionin und Tryptophan werden sämtlich in der pharmazeutischen Industrie verwendet. Glutamin, Valin, Leucin, Isoleucin, Histidin, Arginin, Prolin, Serin und Alanin werden in der pharmazeutischen Industrie und der Kosmetikindustrie verwendet. Threonin, Tryptophan und D-/L-Methionin sind weitverbreitete Futtermittelzusätze (Leuchtenberger, W. (1996) Amino acids - technical production and use, S. 466-502 in Rehm et al., (Hrsg.) Biotechnology Bd. 6, Kapitel 14a, VCH: Weinheim). Man hat entdeckt, daß
- 10 sich diese Aminosäuren außerdem als Vorstufen für die Synthese von synthetischen Aminosäuren und Proteinen, wie N-Acetylcystein, S-Carboxymethyl-L-cystein, (S)-5-Hydroxytryptophan und anderen in Ulmann's Encyclopedia of Industrial Chemistry, Bd. A2, S. 57-97, VCH, Weinheim, 1985 beschriebenen Substanzen eignen.
- 15 Die Biosynthese dieser natürlichen Aminosäuren in Organismen, die sie produzieren können, bspw. Bakterien, ist gut charakterisiert worden (für einen Überblick der bakteriellen Aminosäure-Biosynthese und ihrer Regulation s. Umbarger, H.E. (1978) Ann. Rev. Biochem. 47:533-606). Glutamat wird durch reduktive Aminierung von
- 20 α -Ketoglutarat, einem Zwischenprodukt im Citronensäure-Zyklus, synthetisiert. Glutamin, Prolin und Arginin werden jeweils nacheinander aus Glutamat erzeugt. Die Biosynthese von Serin erfolgt in einem Dreischritt-Verfahren, beginnt mit 3-Phosphoglycerat
- 25 (einem Zwischenprodukt der Glykolyse) und ergibt nach Oxidations-, Transaminierungs- und Hydrolyseschritten diese Aminosäure. Cystein und Glycin werden jeweils aus Serin produziert, und zwar die erstere durch Kondensation von Homocystein mit Serin, und die letztere durch Übertragung des Seitenketten- β -Kohlen-
- 30 stoffatoms auf Tetrahydrofolat in einer durch Serin-Transhydroxymethylase katalysierten Reaktion. Phenylalanin und Tyrosin werden aus den Vorstufen des Glykolyse- und Pentosephosphatweges, Erythrose-4-phosphat und Phosphoenolpyruvat, in einem 9-Schritt-Biosyntheseweg synthetisiert, der sich nur in den letzten beiden
- 35 Schritten nach der Synthese von Präphenat unterscheidet. Tryptophan wird ebenfalls aus diesen beiden Ausgangsmolekülen produziert, jedoch erfolgt dessen Synthese in einem 11-Schritt-Weg. Tyrosin läßt sich in einer durch Phenylalaninhydroxylase katalysierten Reaktion auch aus Phenylalanin herstellen. Alanin, Valin
- 40 und Leucin sind jeweils Biosyntheseprodukte aus Pyruvat, dem Endprodukt der Glykolyse. Aspartat wird aus Oxalacetat, einem Zwischenprodukt des Citratzyklus, gebildet. Asparagin, Methionin, Threonin und Lysin werden jeweils durch Umwandlung von Aspartat produziert. Isoleucin wird aus Threonin gebildet. In einem komplexen 9-Schritt-Weg erfolgt die Bildung von Histidin aus
- 45 5-Phosphoribosyl-1-pyrophosphat, einem aktivierten Zucker.

- Aminosäuren, deren Menge den Proteinbiosynthesebedarf übersteigt, können nicht gespeichert werden, und werden statt dessen abgebaut, so daß Zwischenprodukte für die Haupt-Stoffwechselwege der Zelle bereitgestellt werden (für einen Überblick siehe Stryer, 5 L., Biochemistry, 3. Aufl. Kap. 21 "Amino Acid Degradation and the Urea Cycle"; S. 495-516 (1988)). Die Zelle ist zwar in der Lage, ungewünschte Aminosäuren in nützliche Stoffwechsel-Zwischenprodukte umzuwandeln, jedoch ist die Aminosäureproduktion hinsichtlich der Energie, der Vorstufenmoleküle und der für ihre
- 10 Synthese nötigen Enzyme aufwendig. Es überrascht daher nicht, daß die Aminosäure-Biosynthese durch Feedback-Hemmung reguliert wird, wobei das Vorliegen einer bestimmten Aminosäure ihre eigene Produktion verlangsamt oder ganz beendet (für einen Überblick über Rückkopplungs-Mechanismen bei Aminosäure-Biosynthesewegen, siehe
- 15 Stryer, L., Biochemistry, 3. Aufl., Kap. 24, "Biosynthesis of Amino Acids and Heme", S. 575-600 (1988)). Der Ausstoß einer bestimmten Aminosäure wird daher durch die Menge dieser Aminosäure in der Zelle eingeschränkt.

20 B. *Metabolismus und Verwendungen von Vitaminen, Cofaktoren und Nutrazeutika*

- Vitamine, Cofaktoren und Nutrazeutika umfassen eine weitere Gruppe von Molekülen. Höhere Tiere haben die Fähigkeit verloren,
- 25 diese zu synthetisieren und müssen sie somit aufnehmen, obwohl sie leicht durch andere Organismen, wie Bakterien, synthetisiert werden. Diese Moleküle sind entweder biologisch aktive Moleküle an sich oder Vorstufen von biologisch aktiven Substanzen, die als Elektronenüberträger oder Zwischenprodukte bei einer Reihe von
- 30 Stoffwechselwegen dienen. Diese Verbindungen haben neben ihrem Nährwert auch einen signifikanten industriellen Wert als Farbstoffe, Antioxidantien und Katalysatoren oder andere Verarbeitung-Hilfsstoffe. (Für einen Überblick über die Struktur, Aktivität und die industriellen Anwendungen dieser Verbindungen siehe
- 35 bspw. Ullman's Encyclopedia of Industrial Chemistry, "Vitamins", Bd. A27, S. 443-613, VCH: Weinheim, 1996). Der Begriff "Vitamin" ist im Fachgebiet bekannt und umfaßt Nährstoffe, die von einem Organismus für eine normale Funktion benötigt werden, jedoch nicht von diesem Organismus selbst synthetisiert werden können.
- 40 Die Gruppe der Vitamine kann Cofaktoren und nutrazeutische Verbindungen umfassen. Der Begriff "Cofaktor" umfaßt nicht-proteinartige Verbindungen, die für das Auftreten einer normalen Enzymaktivität nötig sind. Diese Verbindungen können organisch oder anorganisch sein; die erfindungsgemäßen Cofaktor-Moleküle sind
- 45 vorzugsweise organisch. Der Begriff "Nutrazeutikum" umfaßt Nahrungsmittelzusätze, die bei Pflanzen und Tieren, insbesondere dem Menschen, gesundheitsfördernd sind. Beispiele solcher Moleküle

sind Vitamine, Antioxidantien und ebenfalls bestimmte Lipide (z.B. mehrfach ungesättigte Fettsäuren).

- Die Biosynthese dieser Moleküle in Organismen, die zu ihrer Produktion befähigt sind, wie Bakterien, ist umfassend charakterisiert worden (Ullman's Encyclopedia of Industrial Chemistry, "Vitamins", Bd. A27, S. 443-613, VCH: Weinheim, 1996, Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley & Sons; Ong, A.S., Niki, E. und Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for free Radical Research - Asien, abgehalten am 1.-3. Sept. 1994 in Penang, Malaysia, AOCs Press, Champaign, IL X, 374 S).
- 15 Thiamin (Vitamin B₁) wird durch chemisches Kuppeln von Pyrimidin und Thiazol-Einheiten gebildet. Riboflavin (Vitamin B₂) wird aus Guanosin-5'-triphosphat (GTP) und Ribose-5'-phosphat synthetisiert. Riboflavin wiederum wird zur Synthese von Flavinmononukleotid (FMN) und Flavinadenindinukleotid (FAD) eingesetzt. Die Familie von Verbindungen, die gemeinsam als "Vitamin B₆" bezeichnet werden (bspw. Pyridoxin, Pyridoxamin, Pyridoxal-5'-phosphat und das kommerziell verwendete Pyridoxinhydrochlorid), sind alle Derivate der gemeinsamen Struktureinheit 5-Hydroxy-6-methylpyridin. Panthothenat (Pantothersäure, R-(+)-N-(2,4-Dihydroxy-3,3-dimethyl-1-oxobutyl)- β -alanin) kann entweder durch chemische Synthese oder durch Fermentation hergestellt werden. Die letzten Schritte bei der Panthothenat-Biosynthese bestehen aus der ATP-getriebenen Kondensation von β -Alanin und Pantoinsäure. Die für die Biosyntheseschritte für die Umwandlung in Pantoinsäure, in β -Alanin und zur Kondensation in Pantothersäure verantwortlichen Enzyme sind bekannt. Die metabolisch aktive Form von Panthothenat ist Coenzym A, dessen Biosynthese über 5 enzymatische Schritte verläuft. Panthothenat, Pyridoxal-5'-phosphat, Cystein und ATP sind die Vorstufen von Coenzym A. Diese Enzyme katalysieren nicht nur die Bildung von Panthothenat, sondern auch die Produktion von (R)-Pantoinsäure, (R)-Pantolacton, (R)-Panthenol (Provitamin B₅), Pantethein (und seinen Derivaten) und Coenzym A.
- 40 Die Biosynthese von Biotin aus dem Vorstufenmolekül Pimeloyl-CoA in Mikroorganismen ist ausführlich untersucht worden, und man hat mehrere der beteiligten Gene identifiziert. Es hat sich herausgestellt, daß viele der entsprechenden Proteine an der Fe-Cluster-Synthese beteiligt sind und zu der Klasse der nifS-Proteine gehören.
- 45 Die Liponsäure wird von der Octanonsäure abgeleitet und dient als Coenzym beim Energie-Metabolismus, wo sie Bestandteil des Pyruvatdehydrogenasekomplexes und des α -Ketoglutaratdehydro-

genasekomplexes wird. Die Folate sind eine Gruppe von Substanzen, die alle von der Folsäure abgeleitet werden, die wiederum von L-Glutaminsäure, p-Aminobenzoessäure und 6-Methylpterin hergeleitet ist. Die Biosynthese der Folsäure und ihrer Derivate, ausgehend von den Stoffwechselzwischenprodukten Guanosin-5'-triphosphat (GTP), L-Glutaminsäure und p-Aminobenzoessäure ist in bestimmten Mikroorganismen eingehend untersucht worden.

- 10 Corrinioide (wie die Cobalamine und insbesondere Vitamin B₁₂) und die Porphyrine gehören zu einer Gruppe von Chemikalien, die sich durch ein Tetrapyrrol-Ringsystem auszeichnen. Die Biosynthese von Vitamin B₁₂ ist hinreichend komplex, daß sie noch nicht vollständig charakterisiert worden ist, jedoch ist inzwischen ein Großteil der beteiligten Enzyme und Substrate bekannt. Nikotinsäure
- 15 (Nikotinat) und Nikotinamid sind Pyridin-Derivate, die auch als "Niacin" bezeichnet werden. Niacin ist die Vorstufe der wichtigen Coenzyme NAD (Nikotinamidadenindinukleotid) und NADP (Nikotinamidadenindinukleotidphosphat) und ihrer reduzierten Formen.
- 20 Die Produktion dieser Verbindungen im Großmaßstab beruht größtenteils auf zellfreien chemischen Synthesen, obwohl einige dieser Chemikalien, wie Riboflavin, Vitamin B₆, Pantothenat und Biotin, auch durch großangelegte Anzucht von Mikroorganismen produziert worden sind. Nur Vitamin B₁₂ wird aufgrund der Komplexität seiner
- 25 Synthese lediglich durch Fermentation produziert. In-vitro-Verfahren erfordern einen erheblichen Aufwand an Materialien und Zeit und häufig an hohen Kosten.

30 C. Purin-, Pyrimidin-, Nukleosid- und Nukleotid-Metabolismus und Verwendungen

- Gene für den Purin- und Pyrimidin-Stoffwechsel und ihre entsprechenden Proteine sind wichtige Ziele für die Therapie von Tumorerkrankungen und Virusinfektionen. Der Begriff "Purin" oder "Pyrimidin" umfaßt stickstoffhaltige Basen, die Bestandteile der Nukleinsäuren, Coenzyme und Nukleotide sind. Der Begriff "Nukleotid" beinhaltet die grundlegenden Struktureinheiten der Nukleinsäuremoleküle, die eine stickstoffhaltige Base, einen Pentose-Zucker (bei RNA ist der Zucker Ribose, bei DNA ist der Zucker D-
- 40 Desoxyribose) und Phosphorsäure umfassen. Der Begriff "Nukleosid" umfaßt Moleküle, die als Vorstufen von Nukleotiden dienen, die aber im Gegensatz zu den Nukleotiden keine Phosphorsäureeinheit aufweisen. Durch Hemmen der Biosynthese dieser Moleküle oder ihrer Mobilisierung zur Bildung von Nukleinsäuremolekülen ist es
- 45 möglich, die RNA- und DNA-Synthese zu hemmen; wird diese Aktivität zielgerichtet bei Krebszellen gehemmt, läßt sich die Teilungs- und Replikationsfähigkeit von Tumorzellen hemmen. Es gibt

zudem Nukleotide, die keine Nukleinsäuremoleküle bilden, jedoch als Energiespeicher (d.h. AMP) oder als Coenzyme (d.h. FAD und NAD) dienen.

- 5 Mehrere Veröffentlichungen haben die Verwendung dieser Chemikalien für diese medizinischen Indikationen beschrieben, wobei der Purin- und/oder Pyrimidin-Metabolismus beeinflusst wird (bspw. Christopherson, R.I. und Lyons, S.D. (1990) "Potent inhibitors of de novo pyrimidine and purine biosynthesis as chemotherapeutic agents", *Med. Res. Reviews* 10:505-548). Untersuchungen an Enzymen, die am Purin- und Pyrimidin-Metabolismus beteiligt sind, haben sich auf die Entwicklung neuer Medikamente konzentriert, die bspw. als Immunsuppressiva oder Antiproliferantien verwendet werden können (Smith, J.L. (1995) "Enzymes in Nucleotide Synthesis" *Curr. Opin. Struct. Biol.* 5:752-757; (1995) *Biochem. Soc. Transact.* 23:877-902). Die Purin- und Pyrimidinbasen, Nukleoside und Nukleotide haben jedoch auch andere Einsatzmöglichkeiten: als Zwischenprodukte bei der Biosynthese verschiedener Feinchemikalien (z.B. Thiamin, S-Adenosyl-methionin, Folate oder Riboflavin), als Energieträger für die Zelle (bspw. ATP oder GTP) und für Chemikalien selbst, die gewöhnlich als Geschmacksverstärker (bspw. IMP oder GMP) oder für viele medizinische Anwendungen verwendet werden (siehe bspw. Kuninaka, A., (1996) "Nucleotides and Related Compounds in Biotechnology Bd. 6, Rehm et al., Hrsg. VCH: Weinheim, S. 561-612). Enzyme, die am Purin-, Pyrimidin-, Nukleosid- oder Nukleotid-Metabolismus beteiligt sind, dienen auch immer stärker als Ziele, gegen die Chemikalien für den Pflanzenschutz, einschließlich Fungiziden, Herbiziden und Insektiziden, entwickelt werden.
- 30 Der Metabolismus dieser Verbindungen in Bakterien ist charakterisiert worden (für Übersichten siehe bspw. Zalkin, H. und Dixon, J.E. (1992) "De novo purin nucleotide biosynthesis" in *Progress in Nucleic Acids Research and Molecular Biology*, Bd. 42, Academic Press, S. 259-287; und Michal, G. (1999) "Nucleotides and Nucleosides"; Kap. 8 in: *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, Wiley, New York). Der Purin-Metabolismus, das Objekt intensiver Forschung, ist für das normale Funktionieren der Zelle essentiell. Ein gestörter Purin-Metabolismus in höheren Tieren kann schwere Erkrankungen verursachen, bspw. Gicht.
- 40 Die Purinnukleotide werden aus Ribose-5-phosphat über eine Reihe von Schritten über die Zwischenverbindung Inosin-5'-phosphat (IMP) synthetisiert, was zur Produktion von Guanosin-5'-monophosphat (GMP) oder Adenosin-5'-monophosphat (AMP) führt, aus denen sich die als Nukleotide verwendeten Triphosphatformen leicht
- 45 herstellen lassen. Diese Verbindungen werden auch als Energiespeicher verwendet, so daß ihr Abbau Energie für viele verschie-

dene biochemische Prozesse in der Zelle liefert. Die Pyrimidinbiosynthese erfolgt über die Bildung von Uridin-5'-monophosphat (UMP) aus Ribose-5-phosphat. UMP wiederum wird in Cytidin-5'-triphosphat (CTP) umgewandelt. Die Desoxyformen sämtlicher Nukleotide werden in einer Einschnitt-Reduktionsreaktion aus der Diphosphat-Riboseform des Nukleotides zur Diphosphat-Desoxyriboseform des Nukleotides hergestellt. Nach der Phosphorylierung können diese Moleküle an der DNA-Synthese teilnehmen.

10 D. Trehalose-Metabolismus und Verwendungen

Trehalose besteht aus zwei Glucosemolekülen, die über eine α, α -1,1-Bindung miteinander verknüpft sind. Sie wird gewöhnlich in der Nahrungsmittelindustrie als Süßstoff, als Additiv für getrocknete oder gefrorene Nahrungsmittel sowie in Getränken verwendet. Sie wird jedoch auch in der pharmazeutischen Industrie, der Kosmetik- und Biotechnologie-Industrie angewendet (s. bspw. Nishimoto et al., (1998) US-Patent Nr. 5 759 610; Singer, M.A. und Lindquist, S. (1998) Trends Biotech. 16:460-467; Paiva, C.L.A. und Panek, A.D. (1996) Biotech Ann. Rev. 2:293-314; und Shiosaka, M. (1997) J. Japan 172:97-102). Trehalose wird durch Enzyme von vielen Mikroorganismen produziert und auf natürliche Weise in das umgebende Medium abgegeben, aus dem sie durch im Fachgebiet bekannte Verfahren gewonnen werden kann.

25

II. Elemente und Verfahren der Erfindung

Die vorliegende Erfindung beruht zumindest teilweise auf der Entdeckung von neuen Molekülen, die hier als MCP-Nukleinsäure-Moleküle bezeichnet werden. Diese MCP-Nukleinsäuremoleküle eignen sich nicht nur zur Identifikation von *C. glutamicum* oder verwandten Bakterienarten, sondern auch als Marker zur Kartierung des *C. glutamicum*-Genoms und zur Identifizierung von Bakterien, die sich zur Produktion von Feinchemikalien durch z.B. fermentative Verfahren eignen. Die vorliegende Erfindung beruht zumindest teilweise auch auf den MCP-Proteinmolekülen, die von diesen MCP-Nukleinsäuremolekülen codiert werden. Diese MCP-Moleküle können die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren, als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen dienen, Kohlenwasserstoffe abbauen und als Ziele für die Entwicklung therapeutischer pharmazeutischer Verbindungen dienen. In einer Ausführungsform sind die erfindungsgemäßen MCP-Moleküle direkt oder indirekt am Stoffwechselweg einer oder mehrerer Feinchemikalien in *C. glutamicum* beteiligt. In einer bevorzugten Ausführungsform hat die Aktivität der erfindungsgemäßen MCP-Moleküle, an solchen Stoffwechselwegen indirekt oder direkt teilzunehmen,

- eine Auswirkung auf die Produktion einer gewünschten Feinchemikalie durch diesen Mikroorganismus. In einer besonders bevorzugten Ausführungsform ist die Aktivität der erfindungsgemäßen MCP-Moleküle moduliert, so daß die *C. glutamicum*-Stoffwechselwege, an denen die erfindungsgemäßen MCP-Proteine beteiligt sind, hinsichtlich der Effizienz oder des Ausstoßes moduliert werden, was direkt oder indirekt die Produktion oder Effizienz der Produktion einer gewünschten Feinchemikalie durch *C. glutamicum* moduliert.
- 10 Der Begriff "MCP-Protein" oder "MCP-Polypeptid" umfaßt Proteine, die die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren, Kohlenwasserstoffe abbauen, Terpenoide oxidieren, als Zielprotein für Arzneimittelscreening oder -design oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen dienen können.
- 15 Beispiele für MCP-Proteine umfassen solche, die von den in Tabelle 1 und Anhang A aufgeführten MCP-Genen codiert werden. Die Ausdrücke "MCP-Gen" oder "MCP-Nukleinsäuresequenz" umfassen Nukleinsäuresequenzen, die ein MCP-Protein codieren, das aus einem
- 20 codierenden Bereich und entsprechenden untranslatierten 5'- und 3'-Sequenzbereichen besteht. Beispiele für MCP-Gene sind die in Tabelle 1 aufgelisteten. Die Begriffe "Produktion" oder "Produktivität" sind im Fachgebiet bekannt und beinhalten die Konzentration des Fermentationsproduktes (bspw. der gewünschten Feinchemikalie), das innerhalb einer festgelegten Zeitspanne und eines
- 25 festgelegten Fermentationsvolumens gebildet wird (bspw. kg Produkt pro Std. pro l). Der Begriff "Effizienz der Produktion" umfaßt die Zeit, die zur Erzielung einer bestimmten Produktionsmenge nötig ist (bspw. wie lange die Zelle zur Aufrichtung einer
- 30 bestimmten Ausstoßrate einer Feinchemikalie benötigt). Der Begriff "Ausbeute" oder "Produkt/Kohlenstoff-Ausbeute" ist im Fachgebiet bekannt und umfaßt die Effizienz der Umwandlung der Kohlenstoffquelle in das Produkt (d.h. die Feinchemikalie). Dies wird bspw. gewöhnlich ausgedrückt als kg Produkt pro kg Kohlen-
- 35 stoffquelle. Durch Vergrößern der Ausbeute oder Produktion der Verbindung wird die Menge der gewonnenen Moleküle oder der geeigneten gewonnenen Moleküle dieser Verbindung in einer bestimmten Kulturmenge über einen festgelegten Zeitraum erhöht. Die Begriffe "Biosynthese" oder "Biosyntheseweg" sind im Fachgebiet bekannt
- 40 und umfassen die Synthese einer Verbindung, vorzugsweise einer organischen Verbindung, durch eine Zelle aus Zwischenverbindungen, bspw. in einem Mehrschritt- oder stark regulierten Prozeß. Die Begriffe "Abbau" oder "Abbauweg" sind im Fachgebiet bekannt und umfassen die Spaltung einer Verbindung, vorzugsweise einer
- 45 organischen Verbindung, durch eine Zelle in Abbauprodukte (allgemeiner gesagt, kleinere oder weniger komplexe Moleküle), bspw. in einem Mehrschritt- oder stark regulierten Prozeß. Der Begriff

- "Metabolismus" ist im Fachgebiet bekannt und umfaßt die Gesamtheit der biochemischen Reaktionen, die in einem Organismus stattfinden. Der Metabolismus einer bestimmten Verbindung (z.B. der Metabolismus einer Aminosäure, wie Glycin) umfaßt dann sämtliche Biosynthese-, Modifikations- und Abbauege in der Zelle, die diese Verbindung betreffen.

- Die erfindungsgemäßen MCP-Moleküle sind in einer anderen Ausführungsform befähigt, die Produktion eines gewünschten Moleküls, wie einer Feinchemikalie, in einem Mikroorganismus, wie *C. glutamicum*, direkt oder indirekt zu modulieren. Unter Verwendung von Gen-Rekombinationstechniken kann/können ein oder mehrere erfindungsgemäße MCP-Proteine so manipuliert werden, daß seine Funktion moduliert ist. Diese Modulation der Funktion kann zur Modulation der Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien von *C. glutamicum* führen.

- Beispielsweise kann man durch Modifikation der Aktivität eines Proteins, das an der Biosynthese oder am Abbau einer Feinchemikalie beteiligt ist, (d.h. durch Mutagenese des entsprechenden Gens) die Fähigkeit der Zelle, diese Verbindung zu synthetisieren oder abzubauen, direkt modulieren und dadurch die Ausbeute und/oder Effizienz der Produktion der Feinchemikalien modulieren. Ebenso kann man durch Modulation der Aktivität eines Proteins, das einen Feinchemikalien-Stoffwechselweg reguliert, direkt beeinflussen, ob die Produktion der gewünschten Verbindung hoch- oder herunterreguliert wird, was beides die Ausbeute oder Effizienz der Produktion der Feinchemikalie von der Zelle moduliert.

- Die indirekte Modulation der Feinchemikalienproduktion kann auch durch Modifikation der Aktivität eines erfindungsgemäßen Proteins (d.h. durch Mutagenese des entsprechenden Gens) erfolgen, so daß die Fähigkeit der Zelle, zu wachsen und sich zu teilen oder lebensfähig und produktiv zu bleiben, insgesamt erhöht ist. Die Produktion von Feinchemikalien aus *C. glutamicum* wird gewöhnlich durch Fermentationskultur im Großmaßstab dieser Mikroorganismen erzielt, Bedingungen, die für das Wachstum und die Zellteilung häufig suboptimal sind. Durch Verändern eines erfindungsgemäßen Proteins (z.B. eines Streßreaktionsproteins, eines Zellwandproteins oder von Proteinen, die am Stoffwechsel von Verbindungen beteiligt sind, die für das Auftreten von Zellwachstum und -teilung nötig sind, wie Nukleotide und Aminosäuren), so daß ein besseres Überleben, Wachsen und Vermehren in diesen Bedingungen möglich ist, kann es möglich sein, die Anzahl und die Produktivität dieser veränderten *C. glutamicum*-Zellen in Kulturen im Großmaßstab zu steigern, was wiederum zu gesteigerten Ausbeuten und/oder zu gesteigerter Effizienz der Produktion einer oder mehrerer ge-

wünschter Feinchemikalien führen sollte. Ferner sind die Stoffwechselwege einer Zelle notwendigerweise voneinander abhängig und co-reguliert. Durch Ändern der Aktivität irgendeines Stoffwechselwegs in *C. glutamicum* (d.h. durch Ändern der Aktivität eines der erfindungsgemäßen Proteine, das an einem solchen Weg beteiligt ist) ist es möglich, gleichzeitig die Aktivität oder Regulation eines anderen Stoffwechselwegs in diesem Mikroorganismus zu ändern, der direkt an der Synthese oder am Abbau einer Feinchemikalie beteiligt sein kann.

10

Die isolierten erfindungsgemäßen Nukleinsäuresequenzen befinden sich im Genom eines *Corynebacterium glutamicum*-Stammes, der von der American Type Culture Collection unter der Bezeichnung ATCC 13032 erhältlich ist. Die Nukleotidsequenz der isolierten *C. glu-*

15

tamicum-MCP-Nukleinsäuremoleküle und die vorhergesagten Aminosäuresequenzen der *C. glutamicum*-MCP-Proteine sind im Anhang A bzw. B gezeigt. Es wurden Computeranalysen durchgeführt, die viele dieser Nukleotidsequenzen als Sequenzen mit Homologie zu *E. coli*- oder *Bacillus subtilis*-Genen klassifizierten und/oder identifi-

20

zierten.

Die vorliegende Erfindung betrifft auch Proteine, deren Aminosäuresequenz zu einer Aminosäuresequenz in Anhang B im wesentlichen homolog ist. Wie hier verwendet, ist ein Protein, dessen Amino-

25

säuresequenz im wesentlichen homolog zu einer ausgewählten Aminosäuresequenz ist, zumindest zu etwa 50% homolog zu der ausgewählten Aminosäuresequenz, bspw. zur gesamten ausgewählten Aminosäuresequenz. Ein Protein, dessen Aminosäuresequenz zu einer ausgewählten Aminosäuresequenz im wesentlichen homolog ist, kann auch

30

mindestens zu etwa 50-60%, vorzugsweise mindestens zu etwa 60-70%, stärker bevorzugt mindestens zu etwa 70-80%, 80-90% oder 90-95% und am stärksten bevorzugt mindestens zu etwa 96%, 97%, 98%, 99% oder noch homologer zur ausgewählten Aminosäuresequenz sein.

35

Ein erfindungsgemäßes MCP-Protein oder ein biologisch aktiver Abschnitt oder Fragment davon kann die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren, Kohlenwasserstoffe abbauen, Terpe-

40

noide oxidieren, als Ziel für Arzneimittelentwicklung dienen oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen dienen.

In den nachstehenden Unterabschnitten sind verschiedene Aspekte

45

der Erfindung ausführlicher beschrieben:

A. Isolierte Nukleinsäuremoleküle

- Ein Aspekt der Erfindung betrifft isolierte Nukleinsäuremoleküle, die MCP-Moleküle oder biologisch aktive Abschnitte davon codieren, sowie Nukleinsäurefragmente, die zur Verwendung als Hybridisierungssonden oder Primer zur Identifizierung oder Amplifizierung von MCP-codierenden Nukleinsäuren (z.B. MCP-DNA) hinreichen. Diese Nukleinsäuremoleküle können zur Identifikation von *C. glutamicum* oder verwandten Organismen, zur Kartierung des Genoms von *C. glutamicum* oder verwandten Organismen oder zur Identifikation von Mikroorganismen, die zur Produktion von Feinchemikalien, z.B. durch Fermentationsverfahren, geeignet sind, verwendet werden. Der Begriff "Nukleinsäuremolekül", wie hier verwendet, soll DNA-Moleküle (z.B. cDNA oder genomische DNA) und RNA-Moleküle (z.B. mRNA) sowie DNA- oder RNA-Analoga, die mittels Nukleotidanaloga erzeugt werden, umfassen. Dieser Begriff umfaßt zudem die am 3'- und am 5'-Ende des codierenden Genbereichs gelegene untranslatierte Sequenz: mindestens etwa 100 Nukleotide der Sequenz stromaufwärts des 5'-Endes des codierenden Bereichs und mindestens etwa 20 Nukleotide der Sequenz stromabwärts des 3'-Endes des codierenden Bereichs des Gens. Das Nukleinsäuremolekül kann einzelsträngig oder doppelsträngig sein, ist aber vorzugsweise doppelsträngige DNA. Ein "isoliertes" Nukleinsäuremolekül wird von anderen Nukleinsäuremolekülen abgetrennt, die in der natürlichen Quelle der Nukleinsäure zugegen sind. Eine "isolierte" Nukleinsäure hat vorzugsweise keine Sequenzen, die die Nukleinsäure in der genomischen DNA des Organismus, aus dem die Nukleinsäure stammt, natürlicherweise flankieren (bspw. Sequenzen, die sich am 5'- bzw. 3'-Ende der Nukleinsäure befinden). In verschiedenen Ausführungsformen kann bspw. das isolierte MCP-Nukleinsäuremolekül weniger als etwa 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb oder 0,1 kb der Nukleotidsequenzen, die natürlicherweise das Nukleinsäuremolekül in der genomischen DNA der Zelle, aus der die Nukleinsäure stammt (bspw. eine *C. glutamicum*-Zelle) flankieren.
- Ein "isoliertes" Nukleinsäuremolekül, wie ein cDNA-Molekül, kann überdies im wesentlichen frei von anderem zellulären Material oder Kulturmedium sein, wenn es durch rekombinante Techniken hergestellt wird, oder frei von chemischen Vorstufen oder anderen Chemikalien sein, wenn es chemisch synthetisiert wird.
- Ein erfindungsgemäßes Nukleinsäuremolekül, bspw. eine Nukleinsäuremolekül mit einer Nukleotidsequenz aus Anhang A oder ein Abschnitt davon, kann mittels molekularbiologischer Standard-Techniken und der hier bereitgestellten Sequenzinformation isoliert werden. Bspw. kann eine *C. glutamicum*-MCP-cDNA aus einer *C. glutamicum*-Bank isoliert werden, indem eine vollständige Sequenz aus Anhang A oder ein Abschnitt davon als Hybridisierungs-sonde und

20

Standard-Hybridisierungstechniken (wie bspw. beschrieben in Sambrook, J., Fritsch, E.F. und Maniatis, T. Molecular Cloning: A Laboratory Manual. 2. Aufl., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989)

- 5 verwendet werden. Überdies läßt sich ein Nukleinsäuremolekül, umfassend eine vollständige Sequenz aus Anhang A oder einen Abschnitt davon, durch Polymerasekettenreaktion isolieren, wobei die Oligonukleotidprimer, die auf der Basis dieser Sequenz erstellt wurden, verwendet werden (z.B. kann ein Nukleinsäuremolekül, umfassend eine vollständige Sequenz aus Anhang A oder einen Abschnitt davon, durch Polymerasekettenreaktion isoliert werden, indem Oligonukleotidprimer verwendet werden, die auf der Basis dieser gleichen Sequenz aus Anhang A erstellt worden sind). Bspw. läßt sich mRNA aus normalen Endothelzellen isolieren (bspw. durch
- 10 das Guanidiniumthiocyanat-Extraktionsverfahren von Chirgwin et al. (1979) Biochemistry 18:5294-5299), und die cDNA kann mittels reverser Transkriptase (bspw. Moloney-MLV-Reverse-Transkriptase, erhältlich bei Gibco/BRL, Bethesda, MD, oder AMV-Reverse-Transkriptase, erhältlich von Seikagaku America, Inc., St. Petersburg,
- 20 FL) und mittels zufallsgemäßen Polynukleotidprimern oder Oligonukleotidprimern auf der Basis einer der im Anhang A gezeigten Nukleotidsequenzen hergestellt werden. Synthetische Oligonukleotidprimer für die Amplifizierung via Polymerasekettenreaktion lassen sich auf der Basis einer der in Anhang A gezeigten Nukleotid-
- 25 sequenzen erstellen. Eine erfindungsgemäße Nukleinsäure kann mittels cDNA oder alternativ genomischer DNA als Matrize und geeigneten Oligonukleotidprimern gemäß PCR-Standard-Amplifikations-techniken amplifiziert werden. Die so amplifizierte Nukleinsäure kann in einen geeigneten Vektor kloniert werden und durch DNA-Se-
- 30 quenzanalyse charakterisiert werden. Oligonukleotide, die einer MCP-Nukleotidsequenz entsprechen, können ferner durch Standard-Syntheseverfahren, bspw. mit einem automatischen DNA-Synthesegerät, hergestellt werden.
- 35 Bei einer bevorzugten Ausführungsform umfaßt ein erfindungsgemäßes isoliertes Nukleinsäuremolekül eine der in Anhang A aufgeführten Nukleotidsequenzen. Die Sequenzen von Anhang A entsprechen den erfindungsgemäßen MCP-cDNAs aus *Corynebacterium glutamicum*. Diese cDNAs umfassen Sequenzen, die MCP-Proteine (d.h. den
- 40 "codierenden Bereich", der in jeder Sequenz in Anhang A angegeben ist), sowie die 5'- und 3'-untranslatierten Sequenzen, die ebenfalls in Anhang A angegeben sind. Das Nukleinsäuremolekül kann alternativ nur den codierenden Bereich einer der Sequenzen in Anhang A umfassen.

Für die Zwecke dieser Anmeldung hat selbstverständlich jede der in Anhang A angegebenen Sequenzen eine RXA- oder RXN-Identifizierungsnummer, wobei hinter der Bezeichnung "RXA" oder "RXN" 5 Ziffern aufgeführt sind (bspw. RXA00003). Jede dieser Sequenzen umfaßt bis zu drei Abschnitte: einen stromaufwärts gelegenen 5'-Bereich, einen codierenden Bereich, und einen stromabwärts gelegenen Bereich. Jeder dieser drei Bereiche ist durch die gleiche RXA- oder RXN-Bezeichnung gekennzeichnet, um Verwirrung zu vermeiden. Die Bezeichnung "eine der Sequenzen in Anhang A" steht dann für eine der Sequenzen in Anhang A, die sich durch ihre unterschiedlichen RXA- oder RXN-Nummern unterscheiden lassen. Der codierende Bereich jeder Sequenz wird in die entsprechende Aminosäuresequenz translatiert, die in Anhang B angegeben ist. Die Sequenzen in Anhang B werden durch die gleichen RXA- oder RXN-Nummern wie in Anhang A identifiziert, so daß sie sich leicht zuordnen lassen. Bspw. ist die mit RXA00003 bezeichnete Aminosäuresequenz in Anhang B eine Translation des codierenden Bereichs der Nukleotidsequenz des Nukleinsäuremoleküls RXA00003 in Anhang A.

- 20 Bei einer Ausführungsform sollen die erfindungsgemäßen Nukleinsäuremoleküle nicht die in Tabelle 2 zusammengestellten umfassen.

Bei einer weiteren bevorzugten Ausführungsform umfaßt ein erfindungsgemäßes isoliertes Nukleinsäuremolekül ein zu einer der in Anhang A gezeigten Nukleotidsequenzen komplementäres Nukleinsäuremolekül oder einen Abschnitt davon, wobei es sich um ein Nukleinsäuremolekül handelt, das zu einer der in Anhang A gezeigten Nukleotidsequenzen hinreichend komplementär ist, daß es mit einer der in Anhang A angegebenen Sequenzen hybridisieren kann, wodurch ein stabiler Duplex entsteht.

Bei einer weiteren bevorzugten Ausführungsform umfaßt ein erfindungsgemäßes isoliertes Nukleinsäuremolekül eine Nukleotidsequenz, die mindestens etwa 50-60%, vorzugsweise mindestens etwa 60-70%, stärker bevorzugt mindestens etwa 70-80%, 80-90% oder 90-95% und noch stärker bevorzugt mindestens etwa 95%, 96%, 97%, 98%, 99% oder noch homologer zu einer in Anhang A angegebenen Nukleotidsequenz oder einem Abschnitt davon ist. Bei einer weiteren bevorzugten Ausführungsform umfaßt ein erfindungsgemäßes isoliertes Nukleinsäuremolekül eine Nukleotidsequenz, die, bspw. unter stringenten Bedingungen, mit einer der in Anhang A gezeigten Nukleotidsequenzen oder einem Abschnitt davon hybridisiert.

Das erfindungsgemäße Nukleinsäuremolekül kann überdies nur einen Abschnitt des codierenden Bereichs von einer der Sequenzen in Anhang A umfassen, bspw. ein Fragment, das als Sonde oder Primer oder Fragment verwendet werden kann, welches einen biologisch ak-

- tiven Abschnitt eines MCP-Proteins codiert. Die aus der Klonierung der MCP-Gene aus *C. glutamicum* ermittelten Nukleotidsequenzen ermöglichen die Erzeugung von Sonden und Primern, die zur Identifizierung und/oder Klonierung von MCP-Homologa in anderen Zelltypen und Organismen und MCP-Homologa von anderen *Corynebakterien* oder verwandten Arten ausgelegt sind. Die Sonde bzw. der Primer umfaßt gewöhnlich im wesentlichen gereinigtes Oligonukleotid. Das Oligonukleotid umfaßt gewöhnlich einen Nukleotidsequenzbereich, der unter stringenten Bedingungen an mindestens etwa 12, vorzugsweise etwa 25, stärker bevorzugt etwa 40, 50 oder 75 aufeinanderfolgende Nukleotide eines Sense-Stranges von einer der in Anhang A angegebenen Sequenzen, eines Antisense-Stranges von einer der in Anhang A angegebenen Sequenzen oder natürlich vorkommenden Mutanten davon hybridisiert. Primer auf der Basis einer Nukleotidsequenz aus Anhang A können in PCR-Reaktionen zur Klonierung von MCP-Homologa verwendet werden. Sonden auf der Basis der MCP-Nukleotidsequenzen können zum Nachweisen von Transkripten oder genomischen Sequenzen, die das gleiche oder homologe Proteine codieren, verwendet werden. In bevorzugten Ausführungsformen umfaßt die Sonde zudem eine daran gebundene Markierungsgruppe, bspw. ein Radioisotop, eine fluoreszierende Verbindung, ein Enzym oder einen Enzym-Cofaktor. Diese Sonden können als Teil eines diagnostischen Test-Kits zur Identifizierung von Zellen verwendet werden, die ein MCP-Protein mißexpressieren, bspw. durch Messen einer Menge einer MCP-codierenden Nukleinsäure in einer Zellenprobe, bspw. durch Nachweisen der MCP-mRNA-Spiegel oder durch Bestimmen, ob ein genomisches MCP-Gen mutiert oder deletiert ist.
- Bei einer Ausführungsform codiert das erfindungsgemäße Nukleinsäuremolekül ein Protein oder einen Abschnitt davon, der eine Aminosäuresequenz umfaßt, die hinreichend homolog zu einer Aminosäuresequenz von Anhang B ist, daß das Protein oder ein Abschnitt davon die Fähigkeit behält, die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* zu modulieren, Kohlenwasserstoffe abzubauen, Terpene zu oxidieren, als Ziel für Arzneimittelentwicklung zu dienen oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen zu dienen. Wie hier verwendet, betrifft der Begriff "hinreichend homolog" Proteine oder Abschnitte davon, deren Aminosäuresequenzen eine minimale Anzahl identischer oder äquivalenter (bspw. einen Aminosäurerest mit einer ähnlichen Seitenkette wie ein Aminosäurerest in einer der Sequenzen von Anhang B) Aminosäurereste zu einer Aminosäuresequenz aus Anhang B aufweisen, so daß das Protein oder ein Abschnitt davon die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren, Kohlenwasserstoffe

23

abbauen, Terpenoide oxidieren, als Ziel für Arzneimittelentwicklung dienen oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen dienen kann. Beispiele dieser Aktivitäten sind ebenfalls hier beschrieben. Somit trägt die "Funktion eines MCP-Proteins" zur Gesamt-Regulation des Stoffwechselweges einer oder mehrerer Feinchemikalien oder zum Abbau eines Kohlenwasserstoffs oder zur Oxidation eines Terpenoids bei.

Bei einer weiteren Ausführungsform ist das Protein mindestens etwa 50-60%, vorzugsweise mindestens etwa 60-70%, stärker bevorzugt mindestens etwa 70-80%, 80-90%, 90-95% und am stärksten bevorzugt mindestens etwa 96%, 97%, 98%, 99% oder noch homologer zu einer vollständigen Aminosäuresequenz in Anhang B.

- 15 Abschnitte von Proteinen, die von den erfindungsgemäßen MCP-Nukleinsäuremolekülen codiert werden, sind vorzugsweise biologisch aktive Abschnitte von einem der MCP-Proteine. Der Begriff "biologisch aktiver Abschnitt eines MCP-Proteins", wie er hier verwendet wird, soll einen Abschnitt, bspw. eine Domäne/ein Motiv eines MCP-Proteins, umfassen, die/das die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* moduliert, Kohlenwasserstoffe abbaut, Terpenoide oxidiert, als Ziel für Arzneimittelentwicklung oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen dient. Zur Bestimmung, ob ein MCP-Protein oder ein biologisch aktiver Abschnitt davon die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren, Kohlenwasserstoffe abbauen oder Terpenoide oxidieren kann, kann ein Test der enzymatischen Aktivität durchgeführt werden. Diese Testverfahren, wie eingehend beschrieben in Beispiel 8 des Beispielteils, sind dem Fachmann geläufig.

Zusätzliche Nukleinsäurefragmente, die biologisch aktive Abschnitte eines MCP-Proteins codieren, lassen sich durch Isolieren eines Abschnitts von einer der Sequenzen in Anhang B, Exprimieren des codierten Abschnitts des MCP-Proteins oder -Peptides (z.B. durch rekombinante Expression in vitro) und Bestimmen der Aktivität des codierten Abschnittes des MCP-Proteins oder -Peptides herstellen.

40

Die Erfindung umfaßt zudem Nukleinsäuremoleküle, die sich von einer der in Anhang A gezeigten Nukleotidsequenzen (und Abschnitten davon) aufgrund des degenerierten genetischen Codes unterscheiden und somit das gleiche MCP-Protein codieren wie dasjenige, das von den in Anhang A gezeigten Nukleotidsequenzen codiert wird. In einer anderen Ausführungsform hat ein erfindungsgemäßes isoliertes Nukleinsäuremolekül eine Nukleotidsequenz, die ein Protein mit

45

einer in Anhang B gezeigten Aminosäuresequenz codiert. In einer weiteren Ausführungsform codiert das erfindungsgemäße Nukleinsäuremolekül ein *C. glutamicum*-Vollängenprotein, das zu einer Aminosäuresequenz aus Anhang B (codiert von einem in Anhang A gezeigten offenen Leseraster) im wesentlichen homolog ist.

- Zusätzlich zu den in Anhang A gezeigten *C. glutamicum*-MCP-Nukleotidsequenzen, ist dem Fachmann bekannt, daß DNA-Sequenzpolymorphismen, die zu Änderungen in den Aminosäuresequenzen von MCP-Proteinen führen, innerhalb einer Population (bspw. der *C. glutamicum*-Population) existieren können. Diese genetischen Polymorphismen im MCP-Gen können zwischen Individuen innerhalb einer Population aufgrund der natürlichen Variation existieren. Wie hier verwendet, bedeuten die Begriffe "Gen" und "rekombinantes Gen" Nukleinsäuremoleküle mit einem offenen Leseraster, das ein MCP-Protein, vorzugsweise ein *C. glutamicum*-MCP-Protein, codiert. Diese natürlichen Variationen bewirken üblicherweise eine Varianz von 1-5% in der Nukleotidsequenz des MCP-Gens. Sämtliche Nukleotidvariationen und daraus resultierenden Aminosäurepolymorphismen in MCP, die das Ergebnis natürlicher Variation sind und die funktionelle Aktivität von MCP-Proteinen nicht verändern, sollen im Umfang der Erfindung liegen.

- Nukleinsäuremoleküle, die natürlichen Varianten entsprechen, und Nicht-*C. glutamicum*-Homologa der erfindungsgemäßen *C. glutamicum*-MCP-cDNA können aufgrund ihrer Homologie zur hier offenbarten *C. glutamicum*-MCP-Nukleinsäure mit der *C. glutamicum*-cDNA oder einem Abschnitt davon als Hybridisierungs-sonde gemäß Standard-Hybridisierungstechniken unter stringenten Hybridisierungsbedingungen isoliert werden. In einer anderen Ausführungsform ist folglich ein erfindungsgemäßes isoliertes Nukleinsäuremolekül mindestens 15 Nukleotide lang und hybridisiert unter stringenten Bedingungen mit dem Nukleinsäuremolekül, das eine Nukleotidsequenz aus Anhang A umfaßt. In anderen Ausführungsformen ist die Nukleinsäure mindestens 30, 50, 100, 250 Nukleotide lang oder länger. Der Begriff "hybridisiert unter stringenten Bedingungen", wie er hier verwendet wird, soll Hybridisierungs- und Waschbedingungen beschreiben, unter denen Nukleotidsequenzen, die mindestens 60% homolog zueinander sind, gewöhnlich aneinander hybridisiert bleiben. Die Bedingungen sind vorzugsweise derart, daß Sequenzen, die mindestens etwa 65%, stärker bevorzugt mindestens etwa 70% und noch stärker bevorzugt mindestens etwa 75% oder stärker zueinander homolog sind, gewöhnlich aneinander hybridisiert bleiben. Diese stringenten Bedingungen sind dem Fachmann bekannt und lassen sich in Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. finden. Ein bevorzugtes, nicht-einschränkendes Beispiel für stringente Hybridisierungsbedingungen

- gen ist eine Hybridisierung in 6x Natriumchlorid/Natriumcitrat (SSC) bei etwa 45°C, gefolgt von einem oder mehreren Waschschritten in 0,2x SSC, 0,1% SDS bei 50-65°C. Ein erfindungsgemäßes isoliertes Nukleinsäuremolekül, das unter stringenten Bedingungen an
- 5 eine Sequenz aus Anhang A hybridisiert, entspricht vorzugsweise einem natürlich vorkommenden Nukleinsäuremolekül. Wie hier verwendet betrifft ein "natürlich vorkommendes" Nukleinsäuremolekül ein RNA- oder DNA-Molekül mit einer Nukleotidsequenz, die in der Natur vorkommt (bspw. ein natürliches Protein codiert). Bei einer
- 10 Ausführungsform codiert die Nukleinsäure ein natürlich vorkommendes *C. glutamicum*-MCP-Protein.

- Zusätzlich zu natürlich vorkommenden Varianten der MCP-Sequenz, die in der Population existieren können, ist der Fachmann sich
- 15 ebenfalls bewußt darüber, daß Änderungen durch Mutation in eine Nukleotidsequenz von Anhang A eingebracht werden können, was zur Änderung der Aminosäuresequenz des codierten MCP-Proteins führt, ohne daß die Funktionsfähigkeit des MCP-Proteins beeinträchtigt wird. Bspw. lassen sich Nukleotidsusbtitutionen, die an "nicht-
- 20 essentiellen" Aminosäureresten zu Aminosäuresubstitutionen führen, in einer Sequenz von Anhang A herstellen. Ein "nicht-essentieller" Aminosäurerest ist ein Rest, der sich in der Wildtypsequenz von einem der MCP-Proteine (Anhang B) verändern läßt, ohne daß die Aktivität des MCP-Proteins verändert wird, wohingegen ein
- 25 "essentieller" Aminosäurerest für die MCP-Proteinaktivität erforderlich ist. Andere Aminosäurereste jedoch (bspw. nicht-konservierte oder lediglich semikonservierte Aminosäurereste in der Domäne mit MCP-Aktivität) können für die Aktivität nicht essentiell sein und lassen sich somit wahrscheinlich verändern, ohne daß die
- 30 MCP-Aktivität verändert wird.

- Ein weiterer Aspekt der Erfindung betrifft folglich Nukleinsäuremoleküle, die MCP-Proteine codieren, die veränderte Aminosäurereste enthalten, die für die MCP-Aktivität nicht-essentiell sind.
- 35 Diese MCP-Proteine unterscheiden sich in der Aminosäuresequenz von einer Sequenz in Anhang B, behalten aber dennoch mindestens eine der hier beschriebenen MCP-Aktivitäten. Das isolierte Nukleinsäuremolekül umfaßt bei einer Ausführungsform eine Nukleotidsequenz, die ein Protein codiert, das eine Aminosäuresequenz
- 40 umfaßt, die mindestens etwa 50% Homologie zu einer Aminosäuresequenz aus Anhang B aufweist und die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren, Kohlenwasserstoffe abbauen, Terpene oxidieren, als Ziel für Arzneimittelentwicklung dienen oder
- 45 als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen dienen kann. Das von dem Nukleinsäuremolekül codierte Protein weist vorzugsweise mindestens etwa 50-60%, stärker bevorzugt

mindestens etwa 60-70%, noch stärker bevorzugt mindestens etwa 70-80%, 80-90%, 90-95%, und am stärksten bevorzugt mindestens etwa 96%, 97%, 98% oder 99% Homologie zu einer der Sequenzen in Anhang B auf.

5

Zur Bestimmung der prozentualen Homologie von zwei Aminosäuresequenzen (bspw. einer der Sequenzen aus Anhang B und einer mutierten Form davon) oder von zwei Nukleinsäuren werden die Sequenzen für optimale Vergleichszwecke untereinander geschrieben (bspw.

- 10 können Lücken in die Sequenz eines Proteins oder einer Nukleinsäure eingefügt werden, damit ein optimales Alignment mit dem anderen Protein oder der anderen Nukleinsäure erzeugt wird). Die Aminosäurereste oder die Nukleotide werden dann an den entsprechenden Aminosäure- oder Nukleotidpositionen miteinander vergli-
- 15 chen. Wenn eine Position in einer Sequenz (bspw. einer der Sequenzen von Anhang B) vom gleichen Aminosäurerest oder Nukleotid belegt wird, wie an der entsprechenden Stelle in der anderen Sequenz (bspw. einer mutanten Form der aus Anhang B ausgewählten Sequenz), dann sind die Moleküle an dieser Position homolog (d.h.
- 20 der hier verwendete Begriff Aminosäure- oder Nukleinsäure-"Homologie" ist äquivalent zu Aminosäure- oder Nukleinsäure-"Identität"). Die prozentuale Homologie zwischen den beiden Sequenzen ist eine Funktion der Anzahl der identischen Positionen in allen Sequenzen (d.h. % Homologie = Anzahl der identischen Positionen/
- 25 Gesamtanzahl der Positionen x 100).

Ein isoliertes Nukleinsäuremolekül, das ein MCP-Protein codiert, das zu einer Proteinsequenz aus Anhang B homolog ist, kann durch Einbringen von einer oder mehreren Nukleotidsubstitutionen,

- 30 -additionen oder -deletionen in eine Nukleotidsequenz aus Anhang A erzeugt werden, so daß eine oder mehrere Aminosäuresubstitutionen, -additionen oder -deletionen in das codierte Protein eingebracht werden. Die Mutationen können in eine der Sequenzen aus Anhang A durch Standard-Techniken, wie stellengerichtete Mu-
- 35 tagenese und PCR-vermittelte Mutagenese, eingebracht werden. Vorzugsweise werden konservative Aminosäuresubstitutionen an einem oder mehreren der vorhergesagten nicht-essentiellen Aminosäurereste eingeführt. Bei einer "konservativen Aminosäuresubstitution" wird der Aminosäurerest durch einen Aminosäurerest mit einer ähn-
- 40 lichen Seitenkette ausgetauscht. Im Fachgebiet sind Familien von Aminosäureresten mit ähnlichen Seitenketten definiert worden. Diese Familien umfassen Aminosäuren mit basischen Seitenketten (z.B. Lysin, Arginin, Histidin), sauren Seitenketten (z.B. Asparaginsäure, Glutaminsäure), ungeladenen polaren Seitenketten
- 45 (z.B. Glycin, Asparagin, Glutamin, Serin, Threonin, Tyrosin, Cystein), nicht-polaren Seitenketten, (bspw. Alanin, Valin, Leucin, Isoleucin, Prolin, Phenylalanin, Methionin, Tryptophan), beta-

- verzweigten Seitenketten (z.B. Threonin, Valin, Isoleucin) und aromatischen Seitenketten (z.B. Tyrosin, Phenylalanin, Tryptophan, Histidin). Ein vorhergesagter nicht-essentieller Aminosäurerest in einem MCP-Protein wird somit vorzugsweise durch einen
- 5 anderen Aminosäurerest der gleichen Seitenkettenfamilie ausgetauscht. In einer weiteren Ausführungsform können die Mutationen alternativ zufallsgemäß über die gesamte oder einen Teil der MCP-codierenden Sequenz eingebracht werden, bspw. durch Sättigungsmutagenese, und die resultierenden Mutanten können auf eine hier
- 10 beschriebene MCP-Aktivität untersucht werden, um Mutanten zu identifizieren, die eine MCP-Aktivität beibehalten. Nach der Mutagenese von einer der Sequenzen aus Anhang A kann das codierte Protein rekombinant exprimiert werden, und die Aktivität des Proteins kann bspw. mit den hier beschriebenen Tests (siehe Beispiel
- 15 8 des Beispielteils) bestimmt werden.

- Zusätzlich zu den Nukleinsäuremolekülen, die die vorstehend beschriebenen MCP-Proteine codieren, betrifft ein weiterer Aspekt der Erfindung isolierte Nukleinsäuremoleküle, die antisense dazu
- 20 sind. Eine "Antisense"-Nukleinsäure umfaßt eine Nukleotidsequenz, die zu einer "Sense"-Nukleinsäure, welche ein Protein codiert, komplementär ist, bspw. komplementär zum codierenden Strang eines doppelsträngigen cDNA-Moleküls oder komplementär zu einer mRNA-Sequenz. Eine Antisense-Nukleinsäure kann folglich über Wasser-
- 25 stoffbrückenbindungen an eine Sense-Nukleinsäure binden. Die Antisense-Nukleinsäure kann zum gesamten MCP-codierenden Strang oder nur zu einem Abschnitt davon komplementär sein. Bei einer Ausführungsform ist ein Antisense-Nukleinsäuremolekül antisense zu einem "codierenden Bereich" des codierenden Stranges einer Nu-
- 30 kleotidsequenz, die ein MCP-Protein codiert. Der Begriff "codierender Bereich" betrifft den Bereich der Nukleotidsequenz, der Codons umfaßt, die in Aminosäurereste translatiert werden (bspw. umfaßt der gesamte codierende Bereich von SEQ.-ID. RXA00003 die Nukleotide 1 bis 741). Bei einer weiteren Ausführungsform ist das
- 35 Antisense-Nukleinsäuremolekül antisense zu einem "nicht-codierenden Bereich" des codierenden Stranges einer Nukleotidsequenz, die MCP codiert. Der Begriff "nicht-codierender Bereich" betrifft 5'- und 3'-Sequenzen, die den codierenden Bereich flankieren und nicht in Aminosäuren translatiert werden (d.h. die auch als 5'-
- 40 und 3'-untranslatierte Bereiche bezeichnet werden).

- Bei den hier offenbarten Sequenzen des codierenden Stranges, die das MCP codieren (bspw. die Sequenzen aus Anhang A), können die erfindungsgemäßen Antisense-Nukleinsäuren gemäß der Regeln der
- 45 Watson-Crick-Basenpaarung ausgestaltet werden. Das Antisense-Nukleinsäuremolekül kann zum gesamten codierenden Bereich von MCP-mRNA komplementär sein, ist aber stärker bevorzugt ein Oligonu-

- kleotid, das zu lediglich einem Abschnitt des codierenden oder nicht-codierenden Bereichs der MCP-mRNA antisense ist. Das Antisense-Oligonukleotid kann bspw. zum Bereich, der die Translationsstartstelle von MCP-mRNA umgibt, komplementär sein. Ein Antisense-Oligonukleotid kann bspw. etwa 5, 10, 15, 20, 25, 30, 35, 40, 45 oder 50 Nukleotide lang sein. Eine erfindungsgemäße Antisense-Nukleinsäure kann durch chemische Synthese und enzymatische Ligationsreaktionen mittels im Fachgebiet bekannter Verfahren konstruiert werden. Eine Antisense-Nukleinsäure (bspw. ein Antisense-Oligonukleotid) kann chemisch synthetisiert werden, wobei natürlich vorkommende Nukleotide oder verschieden modifizierte Nukleotide verwendet werden, die so gestaltet sind, daß sie die biologische Stabilität der Moleküle erhöhen oder die physikalische Stabilität des Duplexes erhöhen, der zwischen der Antisense- und Sense-Nukleinsäure entstanden ist. Bspw. können Phosphorothioat-Derivate und acridinsubstituierte Nukleotide verwendet werden. Beispiele modifizierter Nukleotide, die zur Erzeugung der Antisense-Nukleinsäure verwendet werden können, sind u.a. 5-Fluoruracil, 5-Bromuracil, 5-Chloruracil, 5-Ioduracil, Hypoxanthin, Xanthin, 4-Acetylcytosin, 5-(Carboxyhydroxymethyl)uracil, 5-Carboxymethylaminomethyl-2-thiouridin, 5-Carboxymethylaminomethyluracil, Dihydrouracil, Beta-D-Galactosylqueosin, Inosin, N6-Iso-pentenyladenin, 1-Methylguanin, 1-Methylinosin, 2,2-Dimethylguanin, 2-Methyladenin, 2-Methylguanin, 3-Methylcytosin, 5-Methylcytosin, N6-Adenin, 7-Methylguanin, 5-Methylaminomethyluracil, 5-Methoxyaminomethyl-2-thiouracil, Beta-D-Mannosylqueosin, 5'-Methoxycarboxymethyluracil, 5-Methoxyuracil, 2-Methylthio-N6-isopentenyladenin, Uracil-5-oxyessigsäure (v), Wybutoxosin, Pseudouracil, Queosin, 2-Thiocytosin, 5-Methyl-2-thiouracil, 2-Thiouracil, 4-Thiouracil, 5-Methyluracil, Uracil-5-oxyessigsäuremethylester, Uracil-5-oxyessigsäure (v), 5-Methyl-2-thiouracil, 3-(3-Amino-3-N-2-carboxypropyl)uracil, (acp3)w und 2,6-Diaminopurin. Die Antisense-Nukleinsäure kann ersatzweise biologisch hergestellt werden, indem ein Expressionsvektor verwendet wird, in den eine Nukleinsäure in Antisense-Richtung subkloniert worden ist (d.h. RNA, die von der eingebrachten Nukleinsäure transkribiert wird, ist zu einer Zielnukleinsäure von Interesse in Antisense-Richtung orientiert, was im nachstehenden Unterabschnitt weiter beschrieben ist).
- Die erfindungsgemäßen Antisense-Nukleinsäuremoleküle werden üblicherweise an eine Zelle verabreicht oder in situ erzeugt, so daß sie mit der zellulären mRNA und/oder der genomischen DNA, die ein MCP-Protein codiert, hybridisieren oder daran binden, so daß die Expression des Proteins, bspw. durch Hemmung der Transkription und/oder Translation, gehemmt wird. Die Hybridisierung kann durch herkömmliche Nukleotid-Komplementarität unter Bildung eines sta-

bilen Duplexes oder bspw. im Fall eines Antisense-Nukleinsäuremoleküls, das DNA-Duplices bindet, durch spezifische Wechselwirkungen in der großen Furche der Doppelhelix erfolgen. Das Antisense-Molekül kann so modifiziert werden, daß es spezifisch an einen

5 Rezeptor oder an ein Antigen bindet, das auf einer ausgewählten Zelloberfläche exprimiert wird, bspw. durch Verknüpfen des Antisense-Nukleinsäuremoleküls mit einem Peptid oder einem Antikörper, das/der an einen Zelloberflächenrezeptor oder Antigen bindet. Das Antisense-Nukleinsäuremolekül kann auch unter Verwendung

10 der hier beschriebenen Vektoren an Zellen verabreicht werden. Zur Erzielung hinreichender intrazellulärer Konzentrationen der Antisense-Moleküle sind Vektorkonstrukte, in denen sich das Antisense-Nukleinsäuremolekül unter der Kontrolle eines prokaryotischen, viralen oder eukaryotischen Promotors befindet, bevorzugt.

- 15 In einer weiteren Ausführungsform ist das erfindungsgemäße Antisense-Nukleinsäuremolekül ein α -anomeres Nukleinsäuremolekül. Ein α -anomeres Nukleinsäuremolekül bildet spezifische doppelsträngige Hybride mit komplementärer RNA, wobei die Stränge im Gegensatz
- 20 zu gewöhnlichen β -Einheiten parallel zueinander verlaufen. (Gaultier et al., (1987) Nucleic Acids Res. 15:6625-6641). Das Antisense-Nukleinsäuremolekül kann zudem ein 2'-O-Methyl-ribonukleotid (Inoue et al., (1987) Nucleic Acids Res. 15:6131-6148) oder ein chimäres RNA-DNA-Analogon (Inoue et al.
- 25 (1987) FEBS Lett. 215:327-330) umfassen.

- In einer weiteren Ausführungsform ist eine erfindungsgemäße Antisense-Nukleinsäure ein Ribozym. Ribozyme sind katalytische RNA-Moleküle mit Ribonukleaseaktivität, die eine einzelsträngige Nu-
- 30 kleinsäure, wie eine mRNA, zu der sie einen komplementären Bereich haben, spalten können. Somit können Ribozyme (z.B. Hammerhead-Ribozyme (beschrieben in Haselhoff und Gerlach (1988) Nature 334:585-591)) zur katalytischen Spaltung von MCP-mRNA-Transkripten verwendet werden, um dadurch die Translation der MCP-mRNA zu
- 35 hemmen. Ein Ribozym mit Spezifität für eine MCP-codierende Nukleinsäure kann auf der Basis der Nukleotidsequenz einer hier offenbarten MCP-cDNA (d.h. RXA00003 in Anhang A) gestaltet werden. Bspw. kann ein Derivat einer Tetrahymena-L-19-IVS-RNA konstruiert werden, wobei die Nukleotidsequenz der aktiven Stelle komplementär zur Nukleotidsequenz ist, die in einer MCP-codierenden mRNA
- 40 gespalten werden soll. S. bspw. Cech et al., US-Patent Nr. 4 987 071 und Cech et al., US-Patent Nr. 5 116 742. Alternativ kann MCP-mRNA zur Selektion einer katalytischen RNA mit spezifischer Ribonukleaseaktivität aus einem Pool von RNA-Molekülen verwendet
- 45 werden. Siehe bspw. Bartel, D., und Szostak, J.W. (1993) Science 261:1411-1418.

Die MCP-Genexpression läßt sich alternativ hemmen, indem Nukleotidsequenzen, die komplementär zum regulatorischen Bereich einer MCP-Nukleotidsequenz sind (bspw. zu einem MCP-Promotor und/oder -Enhancer) so dirigiert werden, daß Triple-Helixstrukturen gebildet werden, die die Transkription eines MCP-Gens in Ziel-Zellen verhindern. Siehe allgemein Helene, C. (1991) Anticancer Drug Res. 6(6) 569-584; Helene, C. et al., (1992) Ann. N. Y. Acad. Sci. 660:27-36; und Maher, L.J. (1992) Bioassays 14(12):807-815.

10 B. Rekombinante Expressionsvektoren und Wirtszellen

Ein weiterer Aspekt der Erfindung betrifft Vektoren, vorzugsweise Expressionsvektoren, die eine Nukleinsäure enthalten, die ein MCP-Protein (oder einen Abschnitt davon) codieren. Wie hier verwendet, betrifft der Begriff "Vektor" ein Nukleinsäuremolekül, das eine andere Nukleinsäure transportieren kann, an welche es gebunden ist. Ein Vektortyp ist ein "Plasmid", was für eine zirkuläre doppelsträngige DNA-Schleife steht, in die zusätzliche DNA-Segmente ligiert werden können. Ein weiterer Vektortyp ist ein viraler Vektor, wobei zusätzliche DNA-Segmente in das virale Genom ligiert werden können. Bestimmte Vektoren können in einer Wirtszelle, in die sie eingebracht worden sind, autonom replizieren (bspw. Bakterienvektoren mit bakteriellem Replikationsursprung und episomale Säugetiervektoren). Andere Vektoren (z.B. nicht-episomale Säugetiervektoren) werden in das Genom einer Wirtszelle beim Einbringen in die Wirtszelle integriert und dadurch zusammen mit dem Wirtsgenom repliziert. Zudem können bestimmte Vektoren die Expression von Genen, mit denen sie funktionsfähig verbunden sind, steuern. Diese Vektoren werden hier als "Expressionsvektoren" bezeichnet. Gewöhnlich haben die Expressionsvektoren, die bei DNA-Rekombinationstechniken verwendet werden können, die Form von Plasmiden. In der vorliegenden Beschreibung können "Plasmid" und "Vektor" austauschbar verwendet werden, da das Plasmid die am häufigsten verwendete Vektorform ist. Die Erfindung soll jedoch andere Expressionsvektorformen, wie virale Vektoren (bspw. replikationsdefiziente Retroviren, Adenoviren und adenoverwandte Viren), die ähnliche Funktionen ausüben, umfassen.

Die erfindungsgemäßen rekombinanten Expressionsvektoren umfassen eine erfindungsgemäße Nukleinsäure in einer Form, die sich zur Expression der Nukleinsäure in einer Wirtszelle eignet, d.h. daß die rekombinanten Expressionsvektoren eine oder mehrere regulatorische Sequenzen, ausgewählt auf der Basis der zur Expression zu verwendenden Wirtszellen, umfassen, die mit der zu exprimierenden Nukleinsäuresequenz funktionsfähig verbunden sind. In einem rekombinanten Expressionsvektor bedeutet "funktionsfähig verbunden", daß die Nukleotidsequenz von Interesse derart an die

- regulatorische(n) Sequenz(en) gebunden ist, daß die Expression der Nukleotidsequenz möglich ist (bspw. in einem in-vitro-Transkriptions-/Translationssystem oder in einer Wirtszelle, wenn der Vektor in die Wirtszelle eingebracht ist). Der Begriff "regulatorische Sequenz" soll Promotoren, Repressorbindungsstellen, Aktivatorbindungsstellen, Enhancerbereiche und andere Expressionskontrollelemente (bspw. Terminatoren, andere Elemente der m-RNA-Sekundärstruktur oder Polyadenylierungssignale) umfassen. Diese regulatorischen Sequenzen sind bspw. beschrieben in Goeddel: Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990). Regulatorische Sequenzen umfassen solche, die die konstitutive Expression einer Nukleotidsequenz in vielen Wirtszelltypen steuern, und solche, die die Expression der Nukleotidsequenz nur in bestimmten Wirtszellen steuern. Der Fachmann ist sich dessen bewußt, daß die Gestaltung eines Expressionsvektors von Faktoren abhängen kann, wie der Wahl der zu transformierenden Wirtszelle, dem gewünschten Ausmaß der Proteinexpression usw. Die erfindungsgemäßen Expressionsvektoren können in die Wirtszellen eingebracht werden, so daß dadurch Proteine oder Peptide, einschließlich der Fusionsproteine oder -peptide, die von den Nukleinsäuren, wie hier beschrieben, codiert werden, hergestellt werden (bspw. MCP-Proteine, mutierte Formen von MCP-Proteinen, Fusionsproteine, usw.).
- Die erfindungsgemäßen rekombinanten Expressionsvektoren können zur Expression von MCP-Proteinen in prokaryotischen oder eukaryotischen Zellen ausgestaltet sein. Bspw. können MCP-Gene in bakteriellen Zellen, wie *C. glutamicum*, Insektenzellen (mit Baculovirus-Expressionsvektoren), Hefe- und anderen Pilzzellen (siehe Romanos, M.A. et al. (1992) "Foreign gene expression in yeast: a review", Yeast 8: 423-488; van den Hondel, C.A.M.J.J. et al. (1991) "Heterologous gene expression in filamentous fungi" in: More Gene Manipulations in Fungi, J.W. Bennet & L.L. Lasure, Hrsg., S. 396-428: Academic Press: San Diego; und van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi. in: Applied Molecular Genetics of Fungi, Peberdy, J.F. et al., Hrsg., S. 1-28, Cambridge University Press: Cambridge), Algenzellen und Zellen vielzelliger Pflanzen (siehe Schmidt, R. und Willmitzer, L. (1988) "High efficiency Agrobacterium tumefaciens-mediated transformation of Arabidopsis thaliana leaf and cotyledon explants" Plant Cell Rep.: 583-586) oder Säugetierzellen exprimiert werden. Geeignete Wirtszellen werden weiter erörtert in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990). Der rekombinante Expressionsvektor kann alternativ,

bspw. mit regulatorischen Sequenzen des T7-Promotors und T7-Polymerase, in vitro transkribiert und translatiert werden.

Die Expression von Proteinen in Prokaryoten erfolgt meist mit Vektoren, die konstitutive oder induzierbare Promotoren enthalten, die die Expression von Fusions- oder Nicht-Fusionsproteinen steuern. Fusionsvektoren steuern eine Reihe von Aminosäuren zu einem darin codierten Protein, gewöhnlich am Aminoterminal des rekombinanten Proteins, bei. Diese Fusionsvektoren haben gewöhnlich drei Aufgaben: 1) die Verstärkung der Expression von rekombinantem Protein; 2) die Erhöhung der Löslichkeit des rekombinanten Proteins; und 3) die Unterstützung der Reinigung des rekombinanten Proteins durch Wirkung als Ligand bei der Affinitätsreinigung. Bei Fusions-Expressionsvektoren wird oft eine proteolytische Spaltstelle an der Verbindungsstelle der Fusionseinheit und des rekombinanten Proteins eingebracht, so daß die Trennung des rekombinanten Proteins von der Fusionseinheit nach der Reinigung des Fusionsproteins möglich ist. Diese Enzyme und ihre entsprechenden Erkennungssequenzen umfassen Faktor Xa, Thrombin und Entero-
erokinase.

Übliche Fusionsexpressionsvektoren umfassen pGEX (Pharmacia Biotech Inc; Smith, D.B. und Johnson, K.S. (1988) Gene 67:31-40), pMAL (New England Biolabs, Beverly, MA) und pRIT 5 (Pharmacia, Piscataway, NJ), bei denen Glutathion-S-Transferase (GST), Maltose E-bindendes Protein bzw. Protein A an das rekombinante Zielprotein fusioniert wird. Bei einer Ausführungsform ist die codierende Sequenz des MCP-Proteins in einen pGEX-Expressionsvektor kloniert, so daß ein Vektor erzeugt wird, der ein Fusionsprotein codiert, umfassend vom N-Terminus zum C-Terminus: GST - Thrombin-Spaltstelle - X-Protein. Das Fusionsprotein kann durch Affinitätschromatographie mittels Glutathion-Agarose-Harz gereinigt werden. Das rekombinante MCP-Protein, das nicht mit GST fusioniert ist, kann durch Spaltung des Fusionsproteins mit Thrombin gewonnen werden.

Beispiele geeigneter induzierbarer Nicht-Fusions-E.-coli-Expressionsvektoren umfassen pTrc (Amann et al., (1988) Gene 69:301-315) und pET 11d (Studier et al. Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Kalifornien (1990) 60-89). Die Zielgenexpression aus dem pTrc-Vektor beruht auf der Transkription durch Wirts-RNA-Polymerase von einem Hybrid-trp-lac-Fusionspromotor. Die Zielgenexpression aus dem pET 11d-Vektor beruht auf der Transkription von einem T7-gn10-lac-Fusions-Promotor, die von einer coexprimierten viralen RNA-Polymerase (T7 gn1) vermittelt wird. Diese virale Polymerase wird von den Wirtsstämmen BL 21 (DE3) oder HMS174

(DE3) von einem residenten λ -Prophagen geliefert, der ein T7 gnl-Gen unter der Transkriptionskontrolle des lacUV 5-Promotors birgt.

- 5 Eine Strategie zur Maximierung der Expression des rekombinanten Proteins ist die Expression des Proteins in einem Wirtsbakterium, dessen Fähigkeit zur proteolytischen Spaltung des rekombinanten Proteins gestört ist (Gottesman, S. Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Kalifornien, (1990) 119-128). Eine weitere Strategie ist die Veränderung der Nukleinsäuresequenz der in einen Expressionsvektor zu inserierenden Nukleinsäure, so daß die einzelnen Codons für jede Aminosäure diejenigen sind, die vorzugsweise in einem zur Expression ausgewählten Bakterium, wie *C. glutamicum*, verwendet werden (Wada et al. (1992) Nucleic Acids Res. 20:2111-2118). Diese Veränderung der erfindungsgemäßen Nukleinsäuresequenzen kann durch Standard-DNA-Synthesetechniken erfolgen.

- Bei einer weiteren Ausführungsform ist der MCP-Protein-Expressionsvektor ein Hefe-Expressionsvektor. Beispiele für Vektoren zur Expression in der Hefe *S. cerevisiae* umfassen pYepSec1 (Baldari et al., (1987) Embo J. 6:229-234), pMFa (Kurjan und Herskowitz (1982) Cell 30:933-943), pJRY88 (Schultz et al. (1987) Gene 54:113-123) sowie pYES2 (Invitrogen Corporation, San Diego, CA).
- 25 Vektoren und Verfahren zur Konstruktion von Vektoren, die sich zur Verwendung in anderen Pilzen, wie filamentösen Pilzen, eignen, umfassen diejenigen, die eingehend beschrieben sind in: van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics of Fungi, J.F. Peberdy et al., Hrsg., S. 1-28, Cambridge University Press: Cambridge.

- Alternativ können die erfindungsgemäßen MCP-Proteine in Insektenzellen unter Verwendung von Baculovirus-Expressionsvektoren exprimiert werden. Baculovirus-Vektoren, die zur Expression von Proteinen in gezüchteten Insektenzellen (bspw. Sf9-Zellen) verfügbar sind, umfassen die pAc-Reihe (Smith et al., (1983) Mol. Cell Biol. 3:2156-2165) und die pVL-Reihe (Lucklow und Summers (1989) Virology 170:31-39).

- 40 In einer weiteren Ausführungsform können die erfindungsgemäßen MCP-Proteine in Zellen einzelliger Pflanzen (wie Algen) oder in Pflanzenzellen höherer Pflanzen (bspw. Spermatophyten, wie Feldfrüchte) exprimiert werden. Beispiele für Pflanzen-Expressionsvektoren umfassen solche, die eingehend beschrieben sind in: Becker, D., Kemper, E., Schell, J. und Masterson, R. (1992) "New plant binary vectors with selectable markers located proximal to

the left border", Plant Mol. Biol. 20:1195-1197; und Bevan, M.W. (1984) "Binary Agrobacterium vectors for plant transformation", Nucl. Acids Res. 12:8711-8721.

- 5 In einer weiteren Ausführungsform wird eine erfindungsgemäße Nukleinsäure in Säugetierzellen mit einem Säugetier-Expressionsvektor exprimiert. Beispiele für Säugetier-Expressionsvektoren umfassen pCDM8 (Seed, B. (1987) Nature 329:840) und pMT2PC (Kaufman et al. (1987) EMBO J. 6:187-195). Bei der Verwendung in Säugetierzellen werden die Kontrollfunktionen des Expressionsvektors oft von viralen regulatorischen Elementen bereitgestellt. Gemeinhin verwendete Promotoren stammen bspw. aus Polyoma, Adenovirus 2, Cytomegalievirus und Simian Virus 40. Weitere geeignete Expressionssysteme für prokaryotische und eukaryotische Zellen
- 10 siehe in Kapitel 16 und 17 von Sambrook, J., Fritsch, E.F. und Maniatis, T., Molecular cloning: A Laboratory Manual, 2. Auflage, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.
- Bei einer weiteren Ausführungsform kann der rekombinante Säugetier-Expressionsvektor die Expression der Nukleinsäure vorzugsweise in einem bestimmten Zelltyp bewirken (bspw. werden gewebespezifische regulatorische Elemente zur Expression der Nukleinsäure verwendet). Gewebespezifische regulatorische Elemente sind im Fachgebiet bekannt. Nicht-einschränkende Beispiele für geeignete gewebespezifische Promotoren umfassen den Albuminpromotor (leberspezifisch; Pinkert et al. (1987) Genes Dev. 1:268-277), lymphoid-spezifische Promotoren (Calame und Eaton (1988) Adv. Immunol. 43:235-275), insbesondere Promotoren von T-Zellrezeptoren (Winoto und Baltimore (1989) EMBO J. 8:729-733) und Immunglobulinen (Banerji et al. (1983) Cell 33:729-740; Queen und Baltimore (1983) Cell 33:741-748), neuronenspezifische Promotoren (bspw. der Neurofilament-Promotor; Byrne und Ruddle (1989) PNAS 86:5473-5477), pankreasspezifische Promotoren (Edlund et al., (1985) Science 230:912-916) und milchdrüsenspezifische Promotoren (bspw. Milchserum-Promotor; US-Patent Nr. 4 873 316 und europäische Patentanmeldungsveröffentlichung Nr. 264 166). Entwicklungsregulierte Promotoren sind ebenfalls umfaßt, bspw. die Maus-hox-Promotoren (Kessel und Gruss (1990) Science 249:374-379) und der α -Fetoprotein-Promotor (Campes und Tilghman (1989) Genes Dev. 3:537-546).
- 20
25
30
35
40

Die Erfindung stellt zudem einen rekombinanten Expressionsvektor bereit, umfassend ein erfindungsgemäßes DNA Molekül, das in Antisense-Richtung in den Expressionsvektor kloniert ist. D.h. daß

45 das DNA-Molekül derart mit einer regulatorischen Sequenz funktionsfähig verbunden ist, daß die Expression (durch Transkription des DNA-Moleküls) eines RNA-Moleküls, das zur MCP-mRNA antisense

- ist, möglich wird. Es können regulatorische Sequenzen ausgewählt werden, die funktionsfähig an eine in Antisense-Richtung klonierte Nukleinsäure gebunden sind und die kontinuierliche Expression des Antisense-RNA-Moleküls in einer Vielzahl von Zelltypen steuern, bspw. können virale Promotoren und/oder Enhancer oder regulatorische Sequenzen ausgewählt werden, die die konstitutive, gewebespezifische oder zelltypspezifische Expression von Antisense-RNA steuern. Der Antisense-Expressionsvektor kann in Form eines rekombinanten Plasmids, Phagemids oder attenuierten Virus vorliegen, in dem Antisense-Nukleinsäuren unter der Kontrolle eines hochwirksamen regulatorischen Bereichs produziert werden, dessen Aktivität durch den Zelltyp bestimmt wird, in den der Vektor eingebracht wird. Für eine Diskussion der Regulation der Genexpression mittels Antisense-Genen siehe Weintraub, H. et al.,
- 15 Antisense-RNA as a molecular tool for genetic analysis, Reviews - Trends in Genetics, Bd. 1(1) 1986.

- Ein weiterer Aspekt der Erfindung betrifft die Wirtszellen, in die ein erfindungsgemäßer rekombinanter Expressionsvektor eingebracht worden ist. Die Begriffe "Wirtszelle" und "rekombinante Wirtszelle" werden hier untereinander austauschbar verwendet. Es ist selbstverständlich, daß diese Begriffe nicht nur eine bestimmte Zielzelle, sondern auch die Nachkommen oder potentiellen Nachkommen dieser Zelle betreffen. Da in aufeinanderfolgenden Generationen aufgrund von Mutation oder Umwelteinflüssen bestimmte Modifikationen auftreten können, sind diese Nachkommen nicht unbedingt mit der Parentalzelle identisch, sind jedoch im Umfang des Begriffs, wie er hier verwendet wird, noch umfaßt.

- 30 Eine Wirtszelle kann eine prokaryotische oder eukaryotische Zelle sein. Bspw. kann ein MCP-Protein in Bakterienzellen, wie *C. glutamicum*, Insektenzellen, Hefe- oder Säugetierzellen (wie Ovarzellen des chinesischen Hamsters (CHO) oder COS-Zellen) exprimiert werden. Andere geeignete Wirtszellen sind dem Fachmann geläufig.
- 35 Mikroorganismen, die mit *Corynebacterium glutamicum* verwandt sind und sich geeignet als Wirtszellen für die erfindungsgemäßen Nukleinsäure- und Proteinmoleküle verwenden lassen, sind in Tabelle 3 aufgeführt.
- 40 Durch herkömmliche Transformations- oder Transfektionsverfahren läßt sich Vektor-DNA in prokaryotische oder eukaryotische Zellen einbringen. Die Begriffe "Transformation" und "Transfektion", "Konjugation" und "Transduktion", wie sie hier verwendet werden, sollen eine Vielzahl von im Stand der Technik bekannten Verfahren zum Einbringen fremder Nukleinsäure (bspw. DNA) in eine Wirtszelle umfassen, einschließlich natürlicher Kompetenz, chemisch vermittelter Übertragung, Calciumphosphat- oder Calciumchlorid-

Copräzipitation, DEAE-Dextran-vermittelter Transfektion, Lipofektion oder Elektroporation. Geeignete Verfahren zur Transformation oder Transfektion von Wirtszellen lassen sich nachlesen in Sambrook et al. (Molecular Cloning: A Laboratory Manual. 2. Aufl., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) und anderen Labor-Handbüchern.

Es ist bekannt, daß für die stabile Transfektion von Säugetierzellen je nach dem verwendeten Expressionsvektor und der verwendeten Transfektionstechnik nur ein kleiner Teil der Zellen die fremde DNA in ihr Genom integrieren kann. Zur Identifizierung und Selektion dieser Integranten wird gewöhnlich ein Gen, das einen selektierbaren Marker (z.B. Resistenz gegen Antibiotika) codiert, zusammen mit dem Gen von Interesse in die Wirtszellen eingebracht. Bevorzugte selektierbare Marker umfassen solche, die die Resistenz gegen Medikamente, wie G418, Hygromycin und Methotrexat, verleihen. Eine Nukleinsäure, die einen selektierbaren Marker codiert, kann in eine Wirtszelle auf dem gleichen Vektor eingebracht werden, wie derjenige, der ein MCP-Protein codiert, oder kann auf einem gesonderten Vektor eingebracht werden. Zellen, die mit der eingebrachten Nukleinsäure stabil transfiziert worden sind, können bspw. durch Medikamentenselektion identifiziert werden (z.B. überleben Zellen, die den selektierbaren Marker integriert haben, wohingegen die anderen Zellen sterben).

Zur Erzeugung eines homolog rekombinierten Mikroorganismus wird ein Vektor hergestellt, der zumindest einen Abschnitt eines MCP-Gens enthält, in den eine Deletion, Addition oder Substitution eingebracht worden ist, um das MCP-Gen zu verändern, bspw. funktionell zu disrumpieren. Dieses MCP-Gen ist vorzugsweise ein *Corynebacterium glutamicum*-MCP-Gen, jedoch kann ein Homologon von einem verwandten Bakterium oder sogar aus einer Säugetier-, Hefe- oder Insektenquelle verwendet werden. Bei einer bevorzugten Ausführungsform ist der Vektor derart ausgestaltet, daß das endogene MCP-Gen bei homologer Rekombination funktionell disrumpiert ist (d.h. nicht länger ein funktionelles Protein codiert; auch als "Knockout"-Vektor bezeichnet). Der Vektor kann alternativ derart ausgestaltet sein, daß das endogene MCP-Gen bei homologer Rekombination mutiert oder anderweitig verändert ist, jedoch noch das funktionelle Protein codiert (z.B. kann der stromaufwärts gelegene regulatorische Bereich derart verändert sein, daß dadurch die Expression des endogenen MCP-Proteins verändert wird.). Der veränderte Abschnitt des MCP-Gens ist im homologen Rekombinationsvektor an seinem 5'- und 3'-Ende von zusätzlicher Nukleinsäure des MCP-Gens flankiert, die eine homologe Rekombination zwischen dem exogenen MCP-Gen, das von dem Vektor getragen wird, und einem

endogenen MCP-Gen in einem Mikroorganismus ermöglicht. Die zusätzliche flankierende MCP-Nukleinsäure ist für eine erfolgreiche homologe Rekombination mit dem endogenen Gen hinreichend lang. Gewöhnlich enthält der Vektor weniger als eine Kilobase flankierende DNA (sowohl am 5'- als auch am 3'-Ende) (siehe z.B. Thomas, K.R. und Capecchi, M.R. (1987) Cell 51:503 für eine Beschreibung von homologen Rekombinationsvektoren). Der Vektor wird in einen Mikroorganismus (z.B. durch Elektroporation) eingebracht, und Zellen, in denen das eingebrachte MCP-Gen mit dem endogenen MCP-Gen homolog rekombiniert ist, werden unter Verwendung im Fachgebiet bekannter Verfahren selektiert.

Bei einer anderen Ausführungsform können rekombinante Mikroorganismen produziert werden, die ausgewählte Systeme enthalten, die eine regulierte Expression des eingebrachten Gens ermöglichen. Der Einschluß eines MCP-Gens in einen Vektor, wodurch es unter die Kontrolle des Lac-Operons gebracht wird, ermöglicht z.B. die Expression des MCP-Gens nur in Gegenwart von IPTG. Diese regulatorischen Systeme sind im Fachgebiet bekannt.

Eine erfindungsgemäße Wirtszelle, wie eine prokaryotische oder eukaryotische Wirtszelle in Kultur, kann zur Produktion (d.h. Expression) eines MCP-Proteins verwendet werden. Die Erfindung stellt zudem Verfahren zur Produktion von MCP-Proteinen unter Verwendung der erfindungsgemäßen Wirtszellen bereit. Bei einer Ausführungsform umfaßt das Verfahren die Anzucht der erfindungsgemäßen Wirtszelle (in die ein rekombinanter Expressionsvektor, der ein MCP-Protein codiert, eingebracht worden ist, oder in deren Genom ein Gen eingebracht worden ist, das ein Wildtyp- oder verändertes MCP-Protein codiert) in einem geeigneten Medium, bis das MCP-Protein produziert worden ist. Das Verfahren umfaßt in einer weiteren Ausführungsform das Isolieren der MCP-Proteine aus dem Medium oder der Wirtszelle.

35 C. Isolierte MCP-Proteine

Ein weiterer Aspekt der Erfindung betrifft isolierte MCP-Proteine und biologisch aktive Abschnitte davon. Ein "isoliertes" oder "gereinigtes" Protein oder biologisch aktiver Abschnitt davon ist im wesentlichen frei von zellulärem Material, wenn es durch DNA-Rekombinationstechniken produziert wird, oder von chemischen Vorstufen oder andern Chemikalien, wenn es chemisch synthetisiert wird. Der Begriff "im wesentlichen frei von zellulärem Material" umfaßt MCP-Proteinpräparationen, in denen das Protein von zellulären Komponenten der Zellen, in denen es natürlich oder rekombinant produziert wird, abgetrennt ist. Bei einer Ausführungsform umfaßt der Ausdruck "im wesentlichen frei von zellulärem Mate-

- rial" MCP-Proteinpräparationen mit weniger als etwa 30% (bezogen auf das Trockengewicht) Nicht-MCP-Protein (hier auch als "kontaminierendes Protein" bezeichnet), stärker bevorzugt weniger als etwa 20%, noch stärker bevorzugt weniger als etwa 10% und am stärksten bevorzugt weniger als etwa 5% Nicht-MCP-Protein. Das MCP-Protein oder ein biologisch aktiver Abschnitt davon enthält nach rekombinanter Produktion ebenfalls vorzugsweise im wesentlichen kein Kulturmedium, d.h. das Kulturmedium macht weniger als etwa 20%, stärker bevorzugt weniger als etwa 10% und am stärksten bevorzugt weniger als etwa 5% des Volumens der Proteinpräparation aus. Der Begriff "im wesentlichen frei von chemischen Vorstufen oder anderen Chemikalien" umfaßt MCP-Proteinpräparationen, in denen das Protein von chemischen Vorstufen oder anderen Chemikalien abgetrennt ist, die an der Synthese des Proteins beteiligt sind.
- 15 Bei einer Ausführungsform umfaßt der Begriff "im wesentlichen frei von chemischen Vorstufen oder anderen Chemikalien" MCP-Proteinpräparationen mit weniger als etwa 30% (bezogen auf das Trockengewicht), stärker bevorzugt weniger als etwa 20%, noch stärker bevorzugt weniger als etwa 10% und am stärksten bevorzugt weniger als etwa 5% chemische Vorstufen oder Nicht-MCP-Chemikalien. In bevorzugten Ausführungsformen weisen die isolierten Proteine oder biologisch aktiven Abschnitte davon keine kontaminierenden Proteine aus dem gleichen Organismus auf, aus dem das MCP-Protein stammt. Diese Proteine werden gewöhnlich durch rekombinante Expression, bspw. eines *C. glutamicum*-MCP-Proteins, in einem Mikroorganismus, wie *C. glutamicum*, hergestellt.

- Ein erfindungsgemäßes isoliertes MCP-Protein oder ein Abschnitt davon kann die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren, Kohlenwasserstoffe abbauen, Terpenoide oxidieren, als Ziel für Arzneimittelentwicklung dienen oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen dienen. In bevorzugten Ausführungsformen umfaßt das Protein oder ein Abschnitt davon eine Aminosäuresequenz, die zu einer Aminosäuresequenz aus Anhang B hinreichend homolog ist, daß das Protein oder der Abschnitt davon die Fähigkeit, die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* zu modulieren, Kohlenwasserstoffe abzubauen, Terpenoide zu oxidieren, als Ziel für Arzneimittelentwicklung zu dienen oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen zu dienen, beibehält. Der Abschnitt des Proteins ist vorzugsweise ein biologisch aktiver Abschnitt, wie hier beschrieben. Bei einer weiteren bevorzugten Ausführungsform hat ein erfindungsgemäßes MCP-Protein eine der in Anhang B gezeigten Aminosäuresequenzen. In einer weiteren bevorzugten Ausführungsform hat das MCP-Protein eine Aminosäuresequenz, die von einer

- Nukleotidsequenz codiert wird, die, bspw. unter stringenten Bedingungen, an eine Nukleotidsequenz aus Anhang A hybridisiert. In noch einer weiteren bevorzugten Ausführungsform hat das MCP-Protein eine Aminosäuresequenz, die von einer Nukleotidsequenz codiert wird und die mindestens etwa 50-60%, vorzugsweise mindestens etwa 60-70%, stärker bevorzugt mindestens etwa 70-80%, 80-90%, 90-95% und noch stärker bevorzugt mindestens etwa 96%, 97%, 98%, 99% oder noch homologer zu einer der Aminosäuresequenzen von Anhang B ist. Die bevorzugten erfindungsgemäßen MCP-Proteine besitzen vorzugsweise ebenfalls mindestens eine der hier beschriebenen MCP-Aktivitäten. Ein bevorzugtes erfindungsgemäßes MCP-Protein umfaßt zum Beispiel eine Aminosäuresequenz, die von einer Nukleotidsequenz codiert wird, die, bspw. unter stringenten Bedingungen, mit einer Nukleotidsequenz von Anhang A hybridisiert, und die die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren, Kohlenwasserstoffe abbauen, Terpenoide oxidieren, als Ziel für Arzneimittelentwicklung dienen oder als Identifikationsmarker für *C. glutamicum* oder verwandte Organismen dienen kann.
- Bei weiteren Ausführungsformen ist das MCP-Protein im wesentlichen homolog zu einer Aminosäuresequenz von Anhang B und behält die funktionelle Aktivität des Proteins mit einer der Sequenzen aus Anhang B und unterscheidet sich dennoch in der Aminosäuresequenz aufgrund von natürlicher Variation oder Mutagenese, wie in Unterabschnitt I oben eingehend beschrieben. In einer weiteren Ausführungsform umfaßt das MCP-Protein folglich eine Aminosäuresequenz, die mindestens etwa 50-60%, vorzugsweise mindestens etwa 60-70%, stärker bevorzugt mindestens etwa 70-80%, 80-90%, 90-95% und am stärksten bevorzugt mindestens etwa 96%, 97%, 98%, 99% oder noch homologer zu einer vollständigen Aminosäuresequenz aus Anhang B ist und die zumindest eine der hier beschriebenen MCP-Aktivitäten aufweist. Bei einer anderen Ausführungsform betrifft die Erfindung ein *C. glutamicum*-Vollängenprotein, das im wesentlichen homolog zu einer vollständigen Aminosäuresequenz aus Anhang B ist.

- Biologisch aktive Abschnitte eines MCP-Proteins umfassen Peptide mit Aminosäuresequenzen, die von der Aminosäuresequenz eines MCP-Proteins hergeleitet sind, bspw. eine in Anhang B gezeigte Aminosäuresequenz oder die Aminosäuresequenz eines Proteins, das zu einem MCP-Protein homolog ist, die weniger Aminosäuren als das Vollängen-MCP-Protein oder das Vollängenprotein aufweisen, das zu einem MCP-Protein homolog ist, und zumindest eine Aktivität eines MCP-Proteins aufweisen. Gewöhnlich umfassen biologisch aktive Abschnitte (Peptide, bspw. Peptide, die bspw. 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 oder mehr Aminosäuren lang sind) eine

Domäne oder ein Motiv mit mindestens einer Aktivität eines MCP-Proteins. Überdies können andere biologisch aktive Abschnitte, in denen andere Bereiche des Proteins deletiert sind, durch rekombinante Techniken hergestellt werden und bezüglich einer oder mehrerer der hier beschriebenen Aktivitäten untersucht werden. Die biologisch aktiven Abschnitte eines MCP-Proteins umfassen vorzugsweise ein oder mehrere ausgewählte Domänen/Motive oder Abschnitte davon mit biologischer Aktivität.

- 10 MCP-Proteine werden vorzugsweise durch DNA-Rekombinationstechniken hergestellt. Bspw wird. ein Nukleinsäuremolekül, das das Protein codiert, in einen Expressionsvektor (wie vorstehend beschrieben) kloniert, der Expressionsvektor wird in eine Wirtszelle (wie vorstehend beschrieben) eingebracht, und das MCP-Protein wird in der Wirtszelle exprimiert. Das MCP-Protein kann dann durch ein geeignetes Reinigungsschema mittels Standard-Protein-Reinigungstechniken aus den Zellen isoliert werden. Alternativ zur rekombinanten Expression kann ein MCP-Protein, -Polypeptid, oder -Peptid mittels Standard-Peptidsynthesetechniken chemisch synthetisiert werden. Überdies kann natives MCP-Protein aus Zellen (bspw. Endothelzellen, Bakterienzellen, Pilzzellen oder anderen Zellen), z.B. mit einem Anti-MCP-Antikörper, isoliert werden, der durch Standardtechniken produziert werden kann, wobei ein erfindungsgemäßes MCP-Protein oder ein Fragment davon verwendet
- 25 wird.

- Die Erfindung stellt auch chimäre MCP-Proteine oder MCP-Fusionsproteine bereit. Wie hier verwendet, umfaßt ein "chimäres MCP-Protein" oder "MCP-Fusionsprotein" ein MCP-Polypeptid, das funktionsfähig an ein Nicht-MCP-Polypeptid gebunden ist. Ein "MCP-Polypeptid" betrifft ein Polypeptid mit einer Aminosäuresequenz, die einem MCP-Protein entspricht, wohingegen ein "Nicht-MCP-Polypeptid" ein Polypeptid mit einer Aminosäuresequenz betrifft, die einem Protein entspricht, das nicht im wesentlichen homolog zum
- 30 MCP-Protein ist, z.B. ein Protein, das sich vom MCP-Protein unterscheidet und vom gleichen oder einem anderen Organismus herrührt. Innerhalb des Fusionsproteins soll der Begriff "funktionsfähig verbunden" bedeuten, daß das MCP-Polypeptid und das Nicht-MCP-Polypeptid im Leseraster miteinander fusioniert sind. Das
 - 40 Nicht-MCP-Polypeptid kann an den N- oder C-Terminus des MCP-Polypeptides gebunden sein. Bei einer Ausführungsform ist das Fusionsprotein bspw. ein GST-MCP-Fusionsprotein, bei dem die MCP-Sequenzen an den C-Terminus der GST-Sequenzen gebunden sind. Diese Fusionsproteine können die Reinigung des rekombinanten MCP-Proteins erleichtern. Bei einer weiteren Ausführungsform ist das Fusionsprotein ein MCP-Protein, das eine heterologe Signalsequenz an seinem N-Terminus aufweist. In bestimmten Wirtszellen (z.B.
 - 45

Säugetier-Wirtszellen) kann die Expression und/oder Sekretion eines MCP-Proteins durch Verwendung einer heterologen Signalsequenz gesteigert werden.

- 5 Ein erfindungsgemäßes chimäres MCP-Protein oder MCP-Fusionsprotein wird durch Standard-DNA-Rekombinationstechniken produziert. DNA-Fragmente, die unterschiedliche Polypeptidsequenzen codieren, werden gemäß herkömmlicher Techniken im Leseraster aneinander ligiert, bspw. durch Einsatz glatter oder überhängender Enden zur
- 10 Ligation, Restriktionsenzymsspaltung zur Bereitstellung geeigneter Enden, Auffüllen kohäsiver Enden, falls erforderlich, Behandlung mit alkalischer Phosphatase, um ungewollte Verknüpfungen zu vermeiden, und enzymatische Ligierung. Bei einer weiteren Ausführungsform kann das Fusionsgen durch herkömmliche Techniken, einschließlich DNA-Syntheseautomaten, synthetisiert werden. Alternativ kann eine PCR-Amplifizierung von Genfragmenten mittels Ankerprimern durchgeführt werden, die komplementäre Überhänge zwischen aufeinanderfolgenden Genfragmenten erzeugen. Diese können anschließend miteinander hybridisiert und reamplifiziert werden, so
- 20 daß eine chimäre Gensequenz erzeugt wird (s. bspw. Current Protocols in Molecular Biology, Ausubel et al., Hrsg., John Wiley & Sons: 1992). Überdies sind viele Expressionsvektoren kommerziell erhältlich, die schon eine Fusionseinheit codieren (bspw. ein GST-Polypeptid). Eine MCP-codierende Nukleinsäure kann in einen
- 25 solchen Expressionsvektor kloniert werden, so daß die Fusionseinheit mit dem MCP-Protein im Leseraster verbunden ist.

- Homologa des MCP-Proteins können durch Mutagenese erzeugt werden, z.B. durch bestimmte Punktmutation oder Verkürzung des MCP-Proteins. Der Begriff "Homologon", wie er hier verwendet wird, be-
- 30 trifft eine variante Form des MCP-Proteins, die als Agonist oder Antagonist der MCP-Protein-Aktivität wirkt. Ein Agonist des MCP-Proteins kann im wesentlichen die gleiche oder einen Teil der biologischen Aktivitäten des MCP-Proteins beibehalten. Ein Anta-
 - 35 gonist des MCP-Proteins kann eine oder mehrere Aktivitäten der natürlich vorkommenden Form des MCP-Proteins bspw. durch kompetitive Bindung an ein stromabwärts oder -aufwärts gelegenes Element eines biochemischen Wegs, der das MCP-Protein enthält, hemmen.
- 40 Bei einer alternativen Ausführungsform können Homologa des MCP-Proteins durch Screening kombinatorischer Banken von Mutanten, bspw. Verkürzungsmutanten, des MCP-Proteins bezüglich MCP-Protein-Agonisten- oder -Antagonisten-Aktivität identifiziert werden. Bei einer Ausführungsform wird eine variierte Bank von
 - 45 MCP-Varianten durch kombinatorische Mutagenese auf Nukleinsäureebene erzeugt und von der variierten Genbank codiert. Eine variierte Bank von MCP-Varianten kann bspw durch enzymatisches

- Ligieren eines Gemisches synthetischer Oligonukleotide in Gensequenzen hergestellt werden, so daß sich ein degenerierter Satz potentieller MCP-Sequenzen als individuelle Polypeptide oder alternativ als Satz größerer Fusionsproteine (z.B. Für Phagen-Display), die diesen Satz von MCP-Sequenzen enthalten, exprimieren läßt. Es gibt eine Vielzahl von Verfahren, die zur Herstellung von Banken potentieller MCP-Homologa aus einer degenerierten Oligonukleotidsequenz verwendet werden können. Die chemische Synthese einer degenerierten Gensequenz kann in einem DNA-Syntheserautomaten durchgeführt werden, und das synthetische Gen kann dann in einen geeigneten Expressionsvektor ligiert werden. Die Verwendung eines degenerierten Gensatzes ermöglicht die Bereitstellung sämtlicher Sequenzen in einem Gemisch, die den gewünschten Satz an potentiellen MCP-Sequenzen codieren. Verfahren zur Synthese degenerierter Oligonukleotide sind im Fachgebiet bekannt (s. bspw. Narang, S.A. (1983) Tetrahedron 39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al., (1984) Science 198:1056; Ike et al. (1983) Nucleic Acids Res. 11:477).
- 20 Zusätzlich können Banken von Fragmenten der MCP-Protein-Codierung verwendet werden, um eine variierte Population von MCP-Fragmenten zum Screening und zur anschließenden Selektion von Homologa eines MCP-Proteins zu erzeugen. Bei einer Ausführungsform kann eine Bank von codierenden Sequenzfragmenten durch Behandeln eines doppelsträngigen PCR-Fragmentes einer codierenden MCP-Sequenz mit einer Nuklease unter Bedingungen, unter denen ein Nicking nur etwa einmal pro Molekül erfolgt, Denaturieren der doppelsträngigen DNA, Renaturieren der DNA unter Bildung doppelsträngiger DNA, die Sense-/Antisense-Paare von verschiedenen genickten Produkten umfassen kann, Entfernen einzelsträngiger Abschnitte aus neu gebildeten Duplices durch Behandlung mit S1-Nuclease und Ligieren der resultierenden Fragmentbank in einen Expressionsvektor erzeugt werden. Durch dieses Verfahren kann eine Expressionsbank hergeleitet werden, die N-terminale, C-terminale und interne Fragmente mit verschiedenen Größen des MCP-Proteins codiert.

- Im Fachgebiet sind mehrere Techniken zum Screening von Genprodukten kombinatorischer Banken, die durch Punktmutationen oder Verkürzung hergestellt worden sind, und zum Screening von cDNA-Banken hinsichtlich Genprodukten mit einer ausgewählten Eigenschaft bekannt. Diese Techniken lassen sich an das schnelle Screening der Genbanken anpassen, die durch kombinatorische Mutagenese von MCP-Homologa erzeugt worden sind. Die am häufigsten verwendeten Techniken zum Screening großer Genbanken, die einer Analyse mit hohem Durchsatz unterliegen, umfassen das Klonieren der Genbank in replizierbare Expressionsvektoren, Transformieren geeigneter Zellen mit der resultierenden Vektorenbank und Exprimieren der

kombinatorischen Gene unter Bedingungen, unter denen der Nachweis der gewünschten Aktivität die Isolation des Vektors, der das Gen codiert, dessen Produkt nachgewiesen wurde, erleichtert. Recursive-Ensemble-Mutagenese (REM), eine neue Technik, die die Häufigkeit funktioneller Mutanten in den Banken vergrößert, kann in Kombination mit den Screeningtests verwendet werden, um MCP-Homologa zu identifizieren (Arkin und Yourvan (1992) PNAS 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3):327-331).

10

Bei einer weiteren Ausführungsform können Tests auf Zellenbasis zur Analyse einer variierten MCP-Bank unter Verwendung von im Fachgebiet bekannten Verfahren verwendet werden.

15 D. Erfindungsgemäße Verwendungen und Verfahren

Die hier beschriebenen Nukleinsäuremoleküle, Proteine, Proteinhomologa, Fusionsproteine, Primer, Vektoren und Wirtszellen können in einem oder mehreren nachstehenden Verfahren verwendet werden:

- 20 Identifikation von *C. glutamicum* und verwandten Organismen, Kartierung von Genomen von Organismen, die mit *C. glutamicum* verwandt sind, Identifikation und Lokalisation von *C. glutamicum*-Sequenzen von Interesse, Evolutionsstudien, Bestimmung von MCP-Proteinbereichen, die für die Funktion notwendig sind, Modulation
- 25 der Aktivität eines MCP-Proteins; Modulation der Aktivität eines oder mehrerer Stoffwechselwege und Modulation der zellulären Produktion einer gewünschten Verbindung, wie einer Feinchemikalie. Die erfindungsgemäßen MCP-Nukleinsäuremoleküle haben eine Vielzahl von Verwendungen. Sie können zunächst zur Identifikation ei-
- 30 nes Organismus als *Corynebacterium glutamicum* oder naher Verwandter davon verwendet werden. Sie können zudem zur Identifikation des Vorliegens von *C. glutamicum* oder eines Verwandten davon in einer Mischpopulation von Mikroorganismen verwendet werden. Die Erfindung stellt die Nukleinsäuresequenzen einer Reihe von *C.*
- 35 *glutamicum*-Genen bereit. Durch Sondieren der extrahierten genomischen DNA einer Kultur einer einheitlichen oder gemischten Population von Mikroorganismen unter stringenten Bedingungen mit einer Sonde, die einen Bereich eines *C. glutamicum*-Gens überspannt, das für diesen Organismus einzigartig ist, kann man bestimmen, ob
- 40 dieser Organismus zugegen ist. *Corynebacterium glutamicum* selbst ist zwar nicht pathogen, jedoch ist es mit pathogenen Arten, wie *Corynebacterium diphtheriae*, verwandt. Der Nachweis eines solchen Organismus ist von signifikanter klinischer Bedeutung.
- 45 Zum Nachweis des Vorliegens von *C. glutamicum* in einer Probe können im Fachgebiet bekannte Techniken eingesetzt werden. Insbesondere können die Zellen in der Probe zunächst in einer geeigneten

- Flüssigkeit oder auf einem geeigneten festen Kulturmedium gezüchtet werden, um die Anzahl der Zellen in der Kultur zu vergrößern. Diese Zellen werden lysiert, und die gesamte enthaltene DNA wird extrahiert und gegebenenfalls gereinigt, um Zelltrümmer und Proteinmaterial zu entfernen, die die anschließende Analyse stören könnten. Polymerasekettenreaktion oder eine ähnliche, im Fachgebiet bekannte Technik wird durchgeführt (s. einen allgemeinen Überblick über Methodologien, die gewöhnlich zur Nukleinsäuresequenz-Amplifikation verwendet werden in Mullis et al., U.S.-Patent Nr. 4683195, Mullis et al., U.S.-Patent Nr. 4965188 und Innis, M.A., und Gelfand, D.H. (1989) PCR-Protocols, A guide to Methods and Applications, Academic Press, S. 3-12, und (1988) Biotechnology 6:1197, und Internationale Patentanmeldung Nr. WO89/01050), wobei Primer, die für ein erfindungsgemäßes MCP-Nukleinsäuremolekül spezifisch sind, mit der Nukleinsäureprobe inkubiert werden, so daß diese bestimmte MCP-Nukleinsäuresequenz, falls in der Probe vorhanden, amplifiziert wird. Die bestimmte, zu amplifizierende Nukleinsäuresequenz wird auf der Basis ihres ausschließlichen Vorkommens im Genom von *C. glutamicum* und nur einiger nah verwandter Bakterien ausgewählt. Das Vorliegen des gewünschten Amplifikationsproduktes zeigt das Vorliegen von *C. glutamicum* oder eines mit *C. glutamicum* nah verwandten Organismus an.
- 25 Die erfindungsgemäßen Nukleinsäure- und Proteinmoleküle können ferner als Marker für spezifische Bereiche des Genoms dienen. Unter Verwendung von im Fachgebiet bekannten Techniken ist es möglich, die physikalische Lokalisierung der erfindungsgemäßen MCP-Nukleinsäuremoleküle auf dem *C. glutamicum*-Genom nachzuweisen,
- 30 was wiederum zur leichteren Lokalisierung anderer Nukleinsäuremoleküle und Gene auf der Karte verwendet werden kann. Die erfindungsgemäßen Nukleinsäuremoleküle können zudem hinreichend homolog zu den Sequenzen verwandter Arten sein, so daß diese Nukleinsäuremoleküle ebenfalls die Konstruktion einer genomischen Karte
- 35 in solchen Bakterien (z.B. *Brevibacterium lactofermentum*) ermöglichen können.
- Die erfindungsgemäßen Nukleinsäure- und Proteinmoleküle eignen sich nicht nur zum Kartieren des Genoms, sondern auch für funktionelle Studien von *C. glutamicum*-Proteinen. Zur Identifikation des Genombereichs, an den ein bestimmtes *C. glutamicum*-DNA-bindendes Protein bindet, kann das *C. glutamicum*-Genom bspw. gespalten und die Fragmente mit dem DNA-bindenden Protein inkubiert werden. Diejenigen, die das Protein binden, können zusätzlich mit
- 45 den erfindungsgemäßen Nukleinsäuremolekülen, vorzugsweise mit leicht nachweisbaren Markierungen, sondiert werden; die Bindung eines solchen Nukleinsäuremoleküls an das Genomfragment ermög-

licht die Lokalisation des Fragmentes auf der genomischen Karte von *C. glutamicum*, und wenn dies mehrmals mit unterschiedlichen Enzymen durchgeführt wird, erleichtert es eine rasche Bestimmung der Nukleinsäuresequenz, an die das Protein bindet.

5

Die erfindungsgemäßen MCP-Nukleinsäuremoleküle eignen sich ebenfalls für Evolutions- und Proteinstruktur-Untersuchungen. Die Stoffwechselprozesse, an denen die erfindungsgemäßen Moleküle beteiligt sind, werden von einer Vielzahl von prokaryotischen und

- 10 eukaryotischen Zellen ausgenutzt; durch Vergleich der Sequenzen der erfindungsgemäßen Nukleinsäuremoleküle mit solchen, die ähnliche Enzyme aus anderen Organismen codieren, kann der Evolutions-Verwandtschaftsgrad der Organismen bestimmt werden. Entsprechend ermöglicht ein solcher Vergleich die Bestimmung, welche Sequenzbereiche konserviert sind und welche nicht, was bei der Bestimmung solcher Bereiche des Proteins hilfreich sein kann, die für die Enzymfunktion essentiell sind. Dieser Typ der Bestimmung ist für Proteintechnologie-Untersuchungen wertvoll und kann einen Hinweis darauf geben, wieviel Mutagenese das Protein tolerieren
- 15 kann ohne die Funktion zu verlieren.
- 20

Die erfindungsgemäßen MCP-Proteine lassen sich als Marker zur Klassifizierung eines unbekannten Bakteriums als *C. glutamicum* oder zur Identifikation von *C. glutamicum* oder nahe verwandten

- 25 Bakterien in einer Probe verwenden. Unter Verwendung von im Fachgebiet bekannten Techniken können bspw. Zellen in einer Probe gegebenenfalls amplifiziert werden (z.B. durch Züchten in einem geeigneten Medium), um die Probengröße zu erhöhen, und können dann lysiert werden, so daß die darin enthaltenen Proteine freigesetzt
- 30 werden. Diese Probe kann gegebenenfalls gereinigt werden, um Zelltrümmer und Nukleinsäuremoleküle zu entfernen, die die anschließende Analyse stören könnten. Antikörper, die für ein ausgewähltes erfindungsgemäßes MCP-Protein spezifisch sind, können mit der Proteinprobe in einem typischen Western-Test-Format inkubiert werden (s. z.B. Ausubel et al., (1988) Current Protocols in
- 35 Molecular Biology, Wiley: New York), wobei der Antikörper an sein Zielprotein bindet, wenn dieses Protein in der Probe vorliegt. Ein MCP-Protein wird für diesen Testtyp ausgewählt, wenn es für *C. glutamicum* oder *C. glutamicum* und sehr nahe verwandte Bakte-
- 40 rien einzigartig oder fast einzigartig ist. Die Proteine in der Probe werden dann durch Gelelektrophorese aufgetrennt und auf eine geeignete Matrix, wie Nitrocellulose übertragen. Ein geeigneter Zweitantikörper mit einer nachweisbaren Markierung (z.B. chemilumineszierend oder colorimetrisch) wird mit der Matrix inkubiert, gefolgt von stringentem Waschen. Das Vorliegen oder Fehlen der Markierung zeigt das Vorliegen oder Fehlen des Zielpro-
- 45 teins in der Probe an. Ist das Protein zugegen, zeigt dies das

Vorliegen von *C. glutamicum* an. Ein ähnliches Verfahren ermöglicht die klassifizierung eines unbekannten Bakteriums als *C. glutamicum*; wenn eine Reihe für *C. glutamicum* spezifischer Proteine nicht in den Proteinproben nachgewiesen wird, die von dem 5 unbekannten Bakterium präpariert wurden, ist dieses Bakterium wahrscheinlich nicht *C. glutamicum*.

Die genetische Manipulation der erfindungsgemäßen MCP-Nukleinsäuremoleküle kann die Produktion von MCP-Proteinen mit funktionellen 10 Unterschieden zu den Wildtyp-MCP-Proteinen bewirken. Diese Proteine können hinsichtlich ihrer Effizienz oder Aktivität verbessert werden, können in größerer Anzahl als gewöhnlich in der Zelle zugegen sein oder können hinsichtlich ihrer Effizienz oder Aktivität geschwächt sein.

15 Diese Änderungen der Aktivität können direkt die Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien in *C. glutamicum* modulieren. Beispielsweise kann man durch Modifikation der Aktivität eines Proteins, das an der 20 Biosynthese oder am Abbau einer Feinchemikalie beteiligt ist, (d.h. durch Mutagenese des entsprechenden Gens) die Fähigkeit der Zelle, diese Verbindung zu synthetisieren oder abzubauen, direkt modulieren und dadurch die Ausbeute und/oder Effizienz der Produktion der Feinchemikalie modulieren. Ebenso kann man durch Modulation der Aktivität eines Proteins, das einen Feinchemikalien- 25 Stoffwechselweg reguliert, direkt beeinflussen, ob die Produktion der gewünschten Verbindung hoch- oder herunterreguliert wird, was beides die Ausbeute oder Effizienz der Produktion der Feinchemikalie von der Zelle moduliert.

30 Die indirekte Modulation der Feinchemikalienproduktion kann auch durch Modifikation der Aktivität eines erfindungsgemäßen Proteins (d.h. durch Mutagenese des entsprechenden Gens) erfolgen, so daß die Fähigkeit der Zelle, zu wachsen und sich zu teilen oder lebensfähig und produktiv zu bleiben, insgesamt erhöht ist. Die 35 Produktion von Feinchemikalien aus *C. glutamicum* wird gewöhnlich durch Fermentationskultur im Großmaßstab dieser Mikroorganismen erzielt, Bedingungen, die für das Wachstum und die Zellteilung häufig suboptimal sind. Durch Verändern eines erfindungsgemäßen Proteins (z.B. eines Streßreaktionsproteins, eines Zellwandproteins oder von Proteinen, die am Stoffwechsel von Verbindungen 40 beteiligt sind, die für das Auftreten von Zellwachstum und -teilung nötig sind, wie Nukleotide und Aminosäuren), so daß ein besseres Überleben, Wachsen und Vermehren in diesen Bedingungen möglich ist, kann es möglich sein, die Anzahl und die Produktivität 45 dieser veränderten *C. glutamicum*-Zellen in Kultur im Großmaßstab zu steigern, was wiederum zu gesteigerten Ausbeuten und/oder zu

gesteigerter Effizienz der Produktion einer oder mehrerer gewünschter Feinchemikalien führen sollte. Ferner sind die Stoffwechselwege einer Zelle notwendigerweise voneinander abhängig und co-reguliert. Durch Ändern der Aktivität irgendeines Stoffwechselwegs in *C. glutamicum* (d.h. durch Ändern der Aktivität eines der erfindungsgemäßen Proteine, das an einem solchen Weg beteiligt ist) ist es möglich, gleichzeitig die Aktivität oder Regulation eines anderen Stoffwechselwegs in diesem Mikroorganismus zu ändern, der direkt an der Synthese oder am Abbau einer Feinchemikalie beteiligt sein kann.

Diese vorstehend genannten Mutagenesestrategien für MCP-Proteine, die erhöhte Ausbeuten einer Feinchemikalie aus *C. glutamicum* bewirken sollen, sollen nicht einschränkend sein; Variationen dieser Mutagenesestrategien sind dem Fachmann leicht ersichtlich. Unter Verwendung dieser Strategien und einschließlich der hier offenbarten Mechanismen können die erfindungsgemäßen Nukleinsäure- und Proteinmoleküle verwendet werden, um *C. glutamicum*- oder verwandte Bakterienstämme, die mutierte MCP-Nukleinsäure- und Proteinmoleküle exprimieren, zu erzeugen, so daß die Ausbeute, Produktion und/oder Effizienz der Produktion einer gewünschten Verbindung verbessert wird. Die gewünschte Verbindung kann jedes von *C. glutamicum* hergestellte Produkt sein, einschließlich der Endprodukte von Biosynthesewegen und Zwischenprodukte natürlich vorkommender metabolischer Wege sowie Moleküle, die im Metabolismus von *C. glutamicum* nicht natürlich vorkommen, die jedoch von einem erfindungsgemäßen *C. glutamicum*-Stamm produziert werden.

Diese Erfindung wird durch die nachstehenden Beispiele weiter veranschaulicht, die nicht als einschränkend aufgefaßt werden sollen. Die Inhalte sämtlicher, in dieser Patentanmeldung zitierter Literaturstellen, Patentanmeldungen, Patente und veröffentlichter Patentanmeldungen sind hiermit durch Bezugnahme aufgenommen.

Beispiele

Beispiel 1: Präparation der gesamten genomischen DNA aus *Corynebacterium glutamicum* ATCC13032

Eine Kultur von *Corynebacterium glutamicum* (ATCC 13032) wurde über Nacht bei 30°C unter starkem Schütteln in BHI-Medium (Difco) gezüchtet. Die Zellen wurden durch Zentrifugation geerntet, der Überstand wurde verworfen, und die Zellen wurden in 5ml Puffer I (5% des Ursprungsvolumens der Kultur - sämtliche angegebenen Volumina sind für 100 ml Kulturvolumen berechnet) resuspendiert.

- Zusammensetzung von Puffer I: 140,34 g/l Saccharose, 2,46 g/l $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$, 10 ml/l KH_2PO_4 -Lösung (100g/l, mit KOH auf pH-Wert 6,7 eingestellt), 50 ml/l M12-Konzentrat (10 g/l $(\text{NH}_4)_2\text{SO}_4$, 1 g/l NaCl, 2 g/l $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$, 0,2 g/l CaCl_2 , 0,5 g/l Hefe-Extrakt
- 5 (Difco), 10 ml/l Spurenelemente-Mischung (200 mg/l $\text{FeSO}_4 \cdot \text{H}_2\text{O}$, 10 mg/l $\text{ZnSO}_4 \cdot 7 \text{H}_2\text{O}$, 3 mg/l $\text{MnCl}_2 \cdot 4 \text{H}_2\text{O}$, 30 mg/l H_3BO_3 , 20 mg/l $\text{CoCl}_2 \cdot 6 \text{H}_2\text{O}$, 1 mg/l $\text{NiCl}_2 \cdot 6 \text{H}_2\text{O}$, 3 mg/l $\text{Na}_2\text{MoO}_4 \cdot 2 \text{H}_2\text{O}$, 500 mg/l Komplexbildner (EDTA oder Citronensäure), 100 ml/l Vitamingemisch (0,2 ml/l Biotin, 0,2 mg/l Folsäure, 20 mg/l p-Aminobenzoessäure,
- 10 20 mg/l Riboflavin, 40 mg/l Ca-Panthothenat, 140 mg/l Nikotinsäure, 40 mg/l Pyridoxolhydrochlorid, 200 mg/l Myo-Inositol). Lysozym wurde in einer Endkonzentration von 2,5 mg/ml zur Suspension gegeben. Nach etwa 4 Std. Inkubation bei 37°C wurde die Zellwand abgebaut, und die erhaltenen Protoplasten wurden durch Zentrifugation geerntet. Das Pellet wurde einmal mit 5 ml Puffer I
- 15 und einmal mit 5 ml TE-Puffer (10 mM Tris-HCl, 1 mM EDTA, pH-Wert 8) gewaschen. Das Pellet wurde in 4 ml TE-Puffer resuspendiert, und 0,5 ml SDS-Lösung (10%) und 0,5 ml NaCl-Lösung (5 M) wurden zugegeben. Nach Zugabe von Proteinase K in einer Endkonzentration
- 20 von 200 µg/ml wurde die Suspension etwa 18 Std. bei 37°C inkubiert. Die DNA wurde durch Extraktion mit Phenol, Phenol-Chloroform-Isoamylalkohol und Chloroform-Isoamylalkohol mittels Standard-Verfahren gereinigt. Dann wurde die DNA durch Zugabe von 1/50 Volumen 3 M Natriumacetat und 2 Volumina Ethanol, anschließend
- 25 Bender Inkubation für 30 min bei -20°C und 30 min Zentrifugation bei 12000 U/min in einer Hochgeschwindigkeitszentrifuge mit einem SS34-Rotor (Sorvall) gefällt. Die DNA wurde in 1 ml TE-Puffer gelöst, der 20 µg/ml RNase A enthielt, und für mindestens 3 Std. bei 4°C gegen 1000 ml TE-Puffer dialysiert. Während dieser Zeit
- 30 wurde der Puffer 3mal ausgetauscht. Zu Aliquots von 0,4 ml der dialysierten DNA-Lösung wurden 0,4 ml 2 M LiCl und 0,8 ml Ethanol zugegeben. Nach 30 min Inkubation bei -20°C wurde die DNA durch Zentrifugation gesammelt (13000 U/min, Biofuge Fresco, Heraeus, Hanau, Deutschland). Das DNA-Pellet wurde in TE-Puffer gelöst.
- 35 Durch dieses Verfahren hergestellte DNA konnte für alle Zwecke verwendet werden, einschließlich Southern-Blotting oder zur Konstruktion genomischer Banken.

Beispiel 2: Konstruktion genomischer *Corynebacterium glutamicum* (ATCC13032)-Banken in *Escherichia coli*

40

Ausgehend von DNA, die wie in Beispiel 1 beschrieben hergestellt wurde, wurden gemäß bekannter und gut eingeführter Verfahren (siehe bspw. Sambrook, J. et al. (1989) "Molecular Cloning: A

45 Laboratory Manual". Cold Spring Harbor Laboratory Press oder

Ausubel, F.M. et al. (1994) "Current Protocols in Molecular Biology", John Wiley & Sons) Cosmid- und Plasmid-Banken hergestellt.

- Es ließ sich jedes Plasmid oder Cosmid einsetzen. Besondere Verwendung fanden die Plasmide pBR322 (Sutcliffe, J.G. (1979) Proc. Natl Acad. Sci. USA, 75:3737-3741); pACYC177 (Change & Cohen (1978) J. Bacteriol. 134:1141-1156); Plasmide der pBS-Reihe (pBSSK+, pBSSK- und andere; Stratagene, LaJolla, USA) oder Cosmide, wie SuperCos1 (Stratagene, LaJolla, USA) oder Lorist6 (Gibson, T.J. Rosenthal, A., und Waterson, R.H. (1987) Gene 53: 283-286.

Beispiel 3: DNA-Sequenzierung und Computer-Funktionsanalyse

- 15 Genomische Banken, wie in Beispiel 2 beschrieben, wurden zur DNA-Sequenzierung gemäß Standard-Verfahren, insbesondere dem Kettenabbruchverfahren mit ABI377-Sequenziermaschinen (s. z.B. Fleischman, R.D. et al. (1995) "Whole-genome Random Sequencing and Assembly of Haemophilus Influenzae Rd.", Science 269:496-512) verwendet. Die Sequenzierprimer mit den folgenden Nukleotidsequenzen wurden verwendet: 5'-GGAAACAGTATGACCATG-3' oder 5'-GTAAACGACGGCCAGT-3'.

Beispiel 4: In-vivo-Mutagenese

- 25 In vivo-Mutagenese von *Corynebacterium glutamicum* kann durchgeführt werden, indem eine Plasmid- (oder andere Vektor-) DNA durch *E. coli* oder andere Mikroorganismen (z.B. *Bacillus* spp. oder Hefen, wie *Saccharomyces cerevisiae*) geschleust wird, die die Integrität ihrer genetischen Information nicht aufrechterhalten können. Übliche Mutatorstämme weisen Mutationen in den Genen für das DNA-Reparatursystem auf (z.B., mutHLS, mutD, mutT, usw., zum Vergleich siehe Rupp, W.D. (1996) DNA repair mechanisms, in: *Escherichia coli* and *Salmonella*, S. 2277-2294, ASM: Washington). Diese 35 Stämme sind dem Fachmann bekannt. Die Verwendung dieser Stämme ist bspw. in Greener, A. und Callahan, M. (1994) Strategies 7:32-34 veranschaulicht.

Beispiel 5: DNA-Transfer zwischen *Escherichia coli* und *Corynebacterium glutamicum*

40

- Mehrere *Corynebacterium*- und *Brevibacterium*-Arten enthalten endogene Plasmide (wie bspw. pHM1519 oder pBL1) die autonom replizieren (für einen Überblick siehe bspw. Martin, J.F. et al. (1987) Biotechnology 5:137-146). Shuttle-Vektoren für *Escherichia coli* und *Corynebacterium glutamicum* lassen sich leicht mittels Standard-Vektoren für *E. coli* konstruieren (Sambrook, J. et al.,

(1989), "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press oder Ausubel, F.M. et al. (1994) "Current Protocols in Molecular Biology", John Wiley & Sons), denen ein Replikationsursprung für und ein geeigneter Marker aus *Corynebacterium glutamicum* beigegeben wird. Solche Replikationsursprünge werden vorzugsweise von endogenen Plasmiden entnommen, die aus *Corynebacterium*- und *Brevibacterium*-Arten isoliert worden sind. Besondere Verwendung als Transformationsmarker für diese Arten sind Gene für Kanamycin-Resistenz (wie solche, die vom Tn5- oder Tn-903-Transposon stammen) oder für Chloramphenicol (Winnacker, E.L. (1987) "From Genes to Clones - Introduction to Gene Technology, VCH, Weinheim). Es gibt zahlreiche Beispiele in der Literatur für die Herstellung einer großen Vielzahl von Shuttle-Vektoren, die in *E. coli* und *C. glutamicum* replizieren und für verschiedene Zwecke verwendet werden können, einschließlich Gen-Überexpression (siehe bspw. Yoshihama, M. et al. (1985) J. Bacteriol. 162:591-597, Martin, J.F. et al., (1987) Biotechnology, 5:137-146 und Eikmanns, B.J. et al. (1992) Gene 102:93-98).

20

Mittels Standard-Verfahren ist es möglich, ein Gen von Interesse in einen der vorstehend beschriebenen Shuttle-Vektoren zu klonieren und solche Hybrid-Vektoren in *Corynebacterium glutamicum*-Stämme einzubringen. Die Transformation von *C. glutamicum* läßt sich durch Protoplastentransformation (Kastsumata, R. et al., (1984) J. Bacteriol. 159:306-311), Elektroporation (Liebl, E. et al., (1989) FEMS Microbiol. Letters, 53:399-303) und in Fällen, bei denen spezielle Vektoren verwendet werden, auch durch Konjugation erzielen (wie z.B. beschrieben in Schäfer, A., et (1990) J. Bacteriol. 172:1663-1666). Es ist ebenfalls möglich, die Shuttle-Vektoren für *C. glutamicum* auf *E. coli* zu übertragen, indem Plasmid-DNA aus *C. glutamicum* (mittels im Fachgebiet bekannter Standard-Verfahren) präpariert und in *E. coli* transformiert wird. Dieser Transformationsschritt kann mit Standard-Verfahren erfolgen, jedoch wird vorteilhafterweise ein Mcr-defizienter *E. coli*-Stamm verwendet, wie NM522 (Gough & Murray (1983) J. Mol. Biol. 166:1-19).

40

Beispiel 6: Bestimmung der Expression des mutanten Proteins

Die Beobachtungen der Aktivität eines mutierten Proteins in einer transformierten Wirtszelle beruhen auf der Tatsache, daß das mutante Protein auf ähnliche Weise und in ähnlicher Menge expriert wird wie das Wildtyp-Protein. Ein geeignetes Verfahren zur Bestimmung der Transkriptionsmenge des mutanten Gens (ein Anzeichen für die mRNA-Menge, die für die Translation des Genprodukts verfügbar ist) ist die Durchführung eines Northern-Blots (s.

bspw. Ausubel et al., (1988) Current Protocols in Molecular Biology, Wiley: New York), wobei ein Primer, der so ausgestaltet ist, daß er an das Gen von Interesse bindet, mit einer nachweisbaren (gewöhnlich radioaktiven oder chemilumineszierenden) Markierung versehen wird, so daß - wenn die Gesamt-RNA einer Kultur des Organismus extrahiert, auf einem Gel aufgetrennt, auf eine stabile Matrix übertragen und mit dieser Sonde inkubiert wird - die Bindung und die Quantität der Bindung der Sonde das Vorliegen und auch die Menge der mRNA für dieses Gen anzeigt. Diese Information ist ein Hinweis auf das Ausmaß der Transkription des mutanten Gens. Gesamt-Zell-RNA läßt sich durch verschiedene Verfahren aus *Corynebacterium glutamicum* isolieren, die im Fachgebiet bekannt sind, wie in Bormann, E.R. et al., (1992) Mol. Microbiol. 6:317-326 beschrieben.

15

Zur Bestimmung des Vorliegens oder der relativen Menge an Protein, das von dieser mRNA translatiert wird, können Standard-Techniken, wie Western-Blot, eingesetzt werden (s. bspw. Ausubel et al. (1988) "Current Protocols in Molecular Biology", Wiley,

20

New York): Bei diesem Verfahren werden Gesamt-Zellproteine extrahiert, durch Gelelektrophorese aufgetrennt, auf eine Matrix, wie Nitrocellulose, übertragen und mit einer Sonde, wie einem Antikörper, die an das gewünschte Protein spezifisch bindet, inkubiert. Diese Sonde ist gewöhnlich mit einer chemilumineszierenden oder colorimetrischen Markierung versehen, die sich leicht nachweisen läßt. Das Vorliegen und die beobachtete Menge an Markierung zeigt das Vorliegen und die Menge des gesuchten Mutantenproteins in der Zelle an.

25

30 Beispiel 7: Wachstum von genetisch verändertem *Corynebacterium glutamicum* - Medien und Anzuchtbedingungen

Genetisch veränderte *Corynebakterien* werden in synthetischen oder natürlichen Wachstumsmedien gezüchtet. Eine Anzahl unterschiedlicher Wachstumsmedien für *Corynebakterien* sind bekannt und leicht erhältlich (Lieb et al. (1989) Appl. Microbiol. Biotechnol. 32:205-210; von der Osten et al. (1998) Biotechnology Letters 11:11-16; Patent DE 4 120 867; Liebl (1992) "The Genus *Corynebacterium*", in: The Prokaryotes, Bd. II, Balows, A., et al., Hrsg. Springer-Verlag). Diese Medien bestehen aus einer oder mehreren Kohlenstoffquellen, Stickstoffquellen, anorganischen Salzen, Vitaminen und Spurenelementen. Bevorzugte Kohlenstoffquellen sind Zucker, wie Mono-, Di- oder Polysaccharide. Sehr gute Kohlenstoffquellen sind bspw. Glucose, Fructose, Mannose, Galactose, Ribose, Sorbose, Ribulose, Lactose, Maltose, Saccharose, Raffinose, Stärke oder Cellulose. Man kann Zucker auch über komplexe Verbindungen, wie Melassen, oder andere Nebenprodukte

35

40

45

- der Zucker-Raffinierung zu den Medien geben. Es kann auch vorteilhaft sein, Gemische verschiedener Kohlenstoffquellen zuzugeben. Andere mögliche Kohlenstoffquellen sind Alkohole und organische Säuren, wie Methanol, Ethanol, Essigsäure oder Milchsäure.
- 5 Stickstoffquellen sind gewöhnlich organische oder anorganische Stickstoffverbindungen oder Materialien, die diese Verbindungen enthalten. Beispielhafte Stickstoffquellen umfassen Ammoniak-Gas oder Ammoniumsalze, wie NH_4Cl oder $(\text{NH}_4)_2\text{SO}_4$, NH_4OH , Nitrate, Harnstoff, Aminosäuren oder komplexe Stickstoffquellen, wie Mais-
- 10 quellwasser, Sojamehl, Sojaprotein, Hefeextrakt, Fleischextrakt und andere.

- Anorganische Salzverbindungen, die in den Medien enthalten sein können, umfassen die Chlorid-, Phosphor- oder Sulfatsalze von
- 15 Calcium, Magnesium, Natrium, Kobalt, Molybdän, Kalium, Mangan, Zink, Kupfer und Eisen. Chelatbildner können zum Medium gegeben werden, um die Metallionen in Lösung zu halten. Besonders geeignete Chelatbildner umfassen Dihydroxyphenole, wie Catechol oder Protocatechuat, oder organische Säuren, wie Citronensäure. Die
- 20 Medien enthalten üblicherweise auch andere Wachstumsfaktoren, wie Vitamine oder Wachstumsförderer, zu denen bspw. Biotin, Riboflavin, Thiamin, Folsäure, Nikotinsäure, Panthothenat und Pyridoxin gehören. Wachstumsfaktoren und Salze stammen häufig von komplexen Medienkomponenten, wie Hefeextrakt, Melassen, Maisquellwasser und
- 25 dergleichen. Die genaue Zusammensetzung der Medienverbindungen hängt stark vom jeweiligen Experiment ab und wird für jeden spezifischen Fall individuell entschieden. Information über die Medienoptimierung ist erhältlich aus dem Lehrbuch "Applied Microbiol. Physiology, A Practical Approach" (Hrsg. P.M. Rhodes, P.F. Stanbury, IRL Press (1997) S. 53-73, ISBN 0 19 963577 3). Wach-
- 30 stumsmedien lassen sich auch von kommerziellen Anbietern beziehen, wie Standard 1 (Merck) oder BHI (Brain heart infusion, DIFCO) und dergleichen.
- 35 Sämtliche Medienkomponenten werden, entweder durch Hitze (20 min bei 1,5 bar und 121°C) oder durch Sterilfiltration, sterilisiert. Die Komponenten können entweder zusammen oder nötigenfalls getrennt sterilisiert werden. Sämtliche Medienkomponenten können zu Beginn der Anzucht zugegen sein oder wahlfrei kontinuierlich oder
- 40 chargenweise hinzugegeben werden.

- Die Anzuchtbedingungen werden für jedes Experiment gesondert definiert. Die Temperatur sollte zwischen 15°C und 45°C liegen und kann während des Experimentes konstant gehalten oder verändert
- 45 werden. Der pH-Wert des Mediums sollte im Bereich von 5 bis 8,5, vorzugsweise um 7,0 liegen, und kann durch Zugabe von Puffern zu den Medien aufrechterhalten werden. Ein beispielhafter Puffer für

diesen Zweck ist ein Kaliumphosphatpuffer. Synthetische Puffer, wie MOPS, HEPES; ACES usw., können alternativ oder gleichzeitig verwendet werden. Der Anzucht-pH-Wert läßt sich während der Anzucht auch durch Zugabe von NaOH oder NH_4OH konstant halten. Werden komplexe Medienkomponenten, wie Hefe-Extrakt, verwendet, sinkt der Bedarf an zusätzlichen Puffern, da viele komplexe Verbindungen eine hohe Pufferkapazität aufweisen. Beim Einsatz eines Fermenters für die Anzucht von Mikroorganismen kann der pH-Wert auch mit gasförmigem Ammoniak reguliert werden.

10

Die Inkubationsdauer liegt gewöhnlich in einem Bereich von mehreren Stunden bis zu mehreren Tagen. Diese Zeit wird so ausgewählt, daß sich die maximale Menge Produkt in der Brühe ansammelt. Die offenbarten Wachstumsexperimente können in einer Vielzahl von Behältern, wie Mikrotiterplatten, Glasröhrchen, Glaskolben oder Glas- oder Metallfermentern unterschiedlicher Größen durchgeführt werden. Zum Screening einer großen Anzahl von Klonen sollten die Mikroorganismen in Mikrotiterplatten, Glasröhrchen oder Schüttelkolben entweder mit oder ohne Schikanen, gezüchtet werden. Vorzugsweise werden 100-ml-Schüttelkolben verwendet, die mit 10% (bezogen auf das Volumen) des erforderlichen Wachstumsmediums gefüllt sind. Die Kolben sollten auf einem Kreisel-schüttler (Amplitude 25 mm) mit einer Geschwindigkeit im Bereich von 100-300 U/min geschüttelt werden. Verdampfungsverluste können durch Aufrechterhalten einer feuchten Atmosphäre verringert werden; alternativ sollte für die Verdampfungsverluste eine mathematische Korrektur durchgeführt werden.

Werden genetisch modifizierte Klone untersucht, sollte auch ein unmodifizierter Kontrollklon oder ein Kontrollklon getestet werden, der das Basisplasmid ohne Insertion enthält. Das Medium wird auf eine OD_{600} von 0,5 - 1,5 angeimpft, wobei Zellen verwendet werden, die auf Agarplatten, wie CM-Platten (10 g/l Glucose, 2,5 g/l NaCl, 2 g/l Harnstoff, 10 g/l Polypepton, 5 g/l Hefeextrakt, 5 g/l Fleischextrakt, 22 g/l Agar pH-Wert 6,8 mit 2 M NaOH), die bei 30°C inkubiert worden sind, gezüchtet wurden. Das Animpfen der Medien erfolgt entweder durch Einbringen einer Kochsalzlösung von *C. glutamicum*-Zellen von CM-Platten oder durch Zugabe einer flüssigen Vorkultur dieses Bakteriums.

40

Beispiel 8: In-vitro-Analyse der Funktion mutanter Proteine

Die Bestimmung der Aktivitäten und kinetischen Parameter von Enzymen ist im Fachgebiet gut bekannt. Experimente zur Bestimmung der Aktivität eines bestimmten veränderten Enzyms müssen an die spezifische Aktivität des Wildtypenzyms angepaßt werden, was innerhalb der Fähigkeiten des Fachmann liegt. Überblicke über En-

- zyme im allgemeinen sowie spezifische Einzelheiten, die die Struktur, Kinetiken, Prinzipien, Verfahren, Anwendungen und Beispiele zur Bestimmung vieler Enzymaktivitäten betreffen, können bspw. in den nachstehenden Literaturstellen gefunden werden: Dixon, M., und Webb, E.C: (1979) Enzymes, Longmans, London; Fersht (1985) Enzyme Structure and Mechanism, Freeman, New York; Walsh (1979) Enzymatic Reaction Mechanisms. Freeman, San Francisco; Price, N.C., Stevens, L. (1982) Fundamentals of Enzymology. Oxford Univ. Press: Oxford; Boyer, P.D: Hrsg. (1983) The Enzymes, 3. Aufl., Academic Press, New York; Bisswanger, H. (1994) Enzymkinetik, 2. Aufl. VCH, Weinheim (ISBN 3527300325); Bergmeyer, H.U., Bergmeyer, J., Graßl, M. Hrsg. (1983-1986) Methods of Enzymatic Analysis, 3. Aufl. Bd. I-XII, Verlag Chemie: Weinheim; und Ullmann's Encyclopedia of Industrial Chemistry (1987) Bd. A9, "Enzymes", VCH, Weinheim, S. 352-363.

- Die Aktivität von Proteinen, die an DNA binden, kann durch viele gut eingeführte Verfahren gemessen werden, wie DNA-Banden-Shift-Assays (die auch als Gelretardations-Assays bezeichnet werden).
- 20 Die Wirkung dieser Proteine auf die Expression anderer Moleküle kann mit Reporter-Gen-Assays (wie in Kolmar, H. et al., (1995) EMBO J. 14:3895-3904 und den darin zitierten Literaturstellen beschrieben) gemessen werden. Reporter-Gen-Testsysteme sind wohl bekannt und für Anwendungen in pro- und eukaryotischen Zellen etabliert, wobei Enzyme, wie beta-Galactosidase, Grün-Fluoreszenz-Protein und mehrere andere verwendet werden.

- Die Bestimmung der Aktivität von Membran-Transportproteinen kann gemäß Techniken, wie sie in Gennis, R.B. (1989) "Pores, Channels and Transporters", in Biomembranes, Molecular Structure and Function, Springer: Heidelberg, S. 85-137; 199-234; und 270-322 beschrieben sind, erfolgen.

- Beispiel 9: Analyse des Einflusses von mutiertem Protein auf die Produktion des gewünschten Produktes

- Die Wirkung der genetischen Modifikation in *C. glutamicum* auf die Produktion einer gewünschten Verbindung (wie einer Aminosäure) kann bestimmt werden, indem die modifizierten Mikroorganismen unter geeigneten Bedingungen (wie den vorstehend beschriebenen) gezüchtet werden und das Medium und/oder die zellulären Komponenten bezüglich der erhöhten Produktion des gewünschten Produktes (d.h. einer Aminosäure) untersucht wird/werden. Solche Analysetechniken sind dem Fachmann wohl bekannt und umfassen Spektroskopie, Dünnschichtchromatographie, Färbeverfahren verschiedener Art, enzymatische und mikrobiologische Verfahren sowie analytische Chromatographie, wie Hochleistungs-Flüssigkeitschromatographie (s. bspw.

- Ullman, Encyclopedia of Industrial Chemistry, Bd. A2, S. 89-90 und S. 443-613, VCH: Weinheim (1985); Fallon, A., et al., (1987) "Applications of HPLC in Biochemistry" in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17; Rehm et al. (1993) 5 Biotechnology, Bd. 3, Kapitel III: "Product recovery and purification", S. 469-714, VCH: Weinheim; Belter, P.A. et al. (1988) Bioseparations: downstream processing for Biotechnology, John Wiley and Sons; Kennedy, J.F. und Cabral, J.M.S. (1992) Recovery processes for biological Materials, John Wiley and Sons; 10 Shaeiwitz, J.A. und Henry, J.D. (1988) Biochemical Separations, in Ullmann's Encyclopedia of Industrial Chemistry, Bd. B3; Kapitel 11, S. 1-27, VCH: Weinheim; und Dechow, F.J. (1989) Separation and purification techniques in biotechnology, Noyes Publications).
- 15 Zusätzlich zur Messung des Fermentationsendproduktes ist es ebenfalls möglich, andere Komponenten der Stoffwechselwege zu analysieren, die zur Produktion der gewünschten Verbindung verwendet werden, wie Zwischen- und Nebenprodukte, um die Gesamt-Effizienz 20 der Produktion der Verbindung zu bestimmen. Die Analyseverfahren umfassen Messungen der Nährstoffmengen im Medium (bspw. Zucker, Kohlenwasserstoffe, Stickstoffquellen, Phosphat und andere Ionen), Messungen der Biomassezusammensetzung und des Wachstums, Analyse der Produktion gemeinsamer Metabolite von Biosynthesewe- 25 gen und Messungen von Gasen, die während der Fermentation erzeugt werden. Standardverfahren für diese Messungen sind in Applied Microbial Physiology; A Practical Approach, P.M. Rhodes und P.F. Stanbury, Hrsg. IRL Press, S. 103-129; 131-163 und 165-192 (ISBN: 0199635773) und den darin angegebenen Literaturstellen beschrie- 30 ben.

Beispiel 10: Reinigung des gewünschten Produktes aus *C. glutamicum*-Kultur

- 35 Die Gewinnung des gewünschten Produktes aus *C. glutamicum*-Zellen oder aus dem Überstand der vorstehend beschriebenen Kultur kann durch verschiedene, im Fachgebiet bekannte Verfahren erfolgen. Wird das gewünschte Produkt von den Zellen nicht sezerniert, können die Zellen aus der Kultur durch langsame Zentrifugation ge- 40 erntet werden, die Zellen können durch Standard-Techniken, wie mechanische Kraft oder Ultraschallbehandlung, lysiert werden. Die Zelltrümmer werden durch Zentrifugation entfernt, und die Überstandsfraktion, die die löslichen Proteine enthält, wird zur weiteren Reinigung der gewünschten Verbindung erhalten. Wird das 45 Produkt von den *C. glutamicum*-Zellen sezerniert, werden die Zel-

len durch langsame Zentrifugation aus der Kultur entfernt, und die Überstandsfraktion wird zur weiteren Reinigung behalten.

Die Überstandsfraktion aus beiden Reinigungsverfahren wird einer Chromatographie mit einem geeigneten Harz unterworfen, wobei das gewünschte Molekül entweder auf dem Chromatographieharz zurückgehalten wird, viele Verunreinigungen in der Probe jedoch nicht, oder die Verunreinigungen auf dem Harz zurückbleiben, die Probe hingegen nicht. Diese Chromatographieschritte können nötigenfalls wiederholt werden, wobei die gleichen oder andere Chromatographieharze verwendet werden. Der Fachmann ist in der Auswahl der geeigneten Chromatographieharze und ihrer wirksamsten Anwendung für ein bestimmtes zu reinigendes Molekül bewandert. Das gereinigte Produkt kann durch Filtration oder Ultrafiltration konzentriert und bei einer Temperatur aufbewahrt werden, bei der die Stabilität des Produktes maximal ist.

Im Fachgebiet sind viele Reinigungsverfahren bekannt, und das vorhergehende Reinigungsverfahren soll nicht einschränkend sein. Diese Reinigungstechniken sind bspw. beschrieben in Bailey, J.E. & Ollis, D.F. Biochemical Engineering Fundamentals, McGraw-Hill: New York (1986).

Die Identität und Reinheit der isolierten Verbindungen kann durch Techniken des Standes der Technik bestimmt werden. Diese umfassen Hochleistungs-Flüssigkeitschromatographie (HPLC), spektroskopische Verfahren, Färbeverfahren, Dünnschichtchromatographie, NIRS, Enzymtest oder mikrobiologische Tests. Diese Analyseverfahren sind zusammengefaßt in: Patek et al. (1994) Appl. Environ. Microbiol. 60:133-140; Malakhova et al. (1996) Biotekhnologiya 11 27-32; und Schmidt et al. (1998) Bioprocess Engineer. 19:67-70. Ulmann's Encyclopedia of Industrial Chemistry (1996) Bd. A27, VCH: Weinheim, S. 89-90, S. 521-540, S. 540-547, S. 559-566, 575-581 und S. 581-587; Michal, G (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley and Sons; Fallon, A. et al. (1987) Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17.

40 Äquivalente

Der Fachmann erkennt oder kann - indem er lediglich Routineverfahren verwendet - viele Äquivalente der erfindungsgemäßen spezifischen Ausführungsformen feststellen. Diese Äquivalente sollen von den nachstehenden Patentansprüchen umfaßt sein.

Tabelle 1: Gene der Patentanmeldung

| | | | | |
|----|----------|--------|-------|-------|
| | RXN00850 | VV0067 | 4148 | 3321 |
| | RXN00869 | VV0127 | 5599 | 4679 |
| 5 | RXN00946 | VV0107 | 3034 | 3807 |
| | RXN01075 | VV0084 | 42045 | 41635 |
| | RXN01134 | VV0077 | 774 | 4 |
| | RXN01140 | VV0077 | 1642 | 710 |
| | RXN01153 | VV0265 | 546 | 4 |
| | RXN01167 | VV0117 | 12777 | 13172 |
| 10 | RXN01246 | VV0104 | 2815 | 3321 |
| | RXN01281 | VV0212 | 4506 | 5267 |
| | RXN01296 | VV0209 | 10462 | 9380 |
| | RXN01324 | VV0082 | 6588 | 6887 |
| | RXN01362 | VV0051 | 27040 | 23387 |
| | RXN01379 | VV0091 | 10518 | 11459 |
| 15 | RXN01400 | VV0126 | 2988 | 1489 |
| | RXN01409 | VV0278 | 5304 | 4483 |
| | RXN01459 | VV0233 | 3311 | 4120 |
| | RXN01471 | VV0019 | 11467 | 10661 |
| | RXN01485 | VV0019 | 26454 | 28505 |
| | RXN01492 | VV0139 | 36004 | 36807 |
| 20 | RXN01574 | VV0009 | 48980 | 47946 |
| | RXN01592 | VV0229 | 14706 | 13405 |
| | RXN01634 | VV0050 | 43466 | 42915 |
| | RXN01694 | VV0139 | 13054 | 13953 |
| | RXN01703 | VV0089 | 7108 | 8220 |

25

30

35

40

45

Tabelle 2

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|--|------------------|---|---|
| A09073 | ppg | Phosphoenolpyruvatcarboxylase | Bachmann, B. et al. "DNA fragment coding for phosphoenolpyruvat carboxylase, recombinant DNA carrying said fragment, strains carrying the recombinant DNA and method for producing L-amino acids using said strains," Patent: EP 0358940-A 3 03/21/90 |
| A45579, A45581, A45583, A45585 A45587 | | Threonindehydratase | Moeckel, B. et al. "Production of L-isoleucine by means of recombinant micro-organisms with deregulated threonine dehydratase," Patent: WO 9519442-A 5 07/20/95 |
| AB003132 | murC; fisQ; fisZ | | Kobayashi, M. et al. "Cloning, sequencing, and characterization of the fisZ gene from coryneform bacteria," <i>Biochem. Biophys. Res. Commun.</i> , 236(2):383-388 (1997) |
| AB015023 | murC; fisQ | | Wachi, M. et al. "A murC gene from Coryneform bacteria," <i>Appl. Microbiol. Biotechnol.</i> , 51(2):223-228 (1999) |
| AB018530 | disR | | Kimura, E. et al. "Molecular cloning of a novel gene, disR, which rescues the detergent sensitivity of a mutant derived from <i>Brevibacterium lactofermentum</i> ," <i>Biosci. Biotechnol. Biochem.</i> , 60(10):1565-1570 (1996) |
| AB018531 | disR1; disR2 | | |
| AB020624 | murI | D-Glutamatracemase | |
| AB023377 | tkt | Transketolase | |
| AB024708 | glbB; gltD | Glutamin-2-oxoglutarataminotransferase große und kleine Untereinheiten | |
| AB025424 | acn | Aconitase | |
| AB027714 | rep | Replikationsprotein | |
| AB027715 | rep; aad | Replikationsprotein; Aminoglycosid- adenyltransferase | |

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|-------------------------|--|---|---|
| AF005242 | argC | N-Acetylglutamat-5-semialdehyd- dehydrogenase | |
| AF005635 | glnA | Glutaminsynthetase | |
| AF030405 | hisF | Cyclase | |
| AF030520 | argG | Argininosuccinatsynthetase | |
| AF031518 | argF | Ornithincarbamolytransferase | |
| AF036932 | aroD | 3-Dehydroquinatedehydratase | |
| AF038548 | pyc | Pyruvatcarboxylase | |
| AF038651 | dciAE; apt; rel | Dipeptid-bindendes Protein; Adenin- phosphoribosyltransferase; GTP- pyrophosphokinase | Wehmeier, L. et al. "The role of the <i>Corynebacterium glutamicum</i> rel gene in (p)ppGpp metabolism," <i>Microbiology</i> , 144:1853-1862 (1998) |
| AF041436 | argR | Arginine-Repressor | |
| AF045998 | impA | Inositolmonophosphatphosphatase | |
| AF048764 | argH | Argininosuccinatlase | |
| AF049897 | argC; argJ; argB; argD; argF; argR; argG; argH | N-Acetylglutamyolphosphatereduktase; Ornithinacetyltransferase; N-Acetyl- glutamatkinase; Acetylornithin-trans- aminase; Ornithin-carbamoyltransferase; Argininrepressor; Argininosuccinatsynthase; Argininosuccinatlase | |
| AF050109 | inhA | Enoyl-acyl-Carrierprotein-Reductase | |
| AF050166 | hisG | ATP-Phosphoribosyltransferase | |
| AF051846 | hisA | Phosphoribosylformi- mino-5-amino-1-phosphoribosyl-4- imidazolecarboxamide isomerase | |
| AF052652 | metA | Homoserin-O-acetyltransferase | Park, S. et al. "Isolation and analysis of metA, a methionine biosynthetic gene encoding homoserine acetyltransferase in <i>Corynebacterium glutamicum</i> ," <i>Mol. Cells</i> , 8(3):286-294 (1998) |
| AF053071 | aroB | Dehydrochinatsynthetase | |
| AF060558 | hisH | Glutaminamidotransferase | |

| GenBank TM Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|--------------------------------------|--------------------------------|---|--|
| AF086704 | hisE | Phosphoribosyl-ATP-pyrophospho- hydrolase | |
| AF114233 | aroA | 5-Enolpyruvylshikimat-3-phosphatsynthase | |
| AF116184 | panD | L-Aspartat- α -decarboxylase-Vorstufe | Dusch, N. et al. "Expression of the <i>Corynebacterium glutamicum</i> panD gene encoding L-aspartate- α -decarboxylase leads to pantothenate overproduction in <i>Escherichia coli</i> ," <i>Appl. Environ. Microbiol.</i> , 65(4):1530-1539 (1999) |
| AF124518 | aroD; aroE | 3-Dehydrochinase; Shikimatdehydrogenase | |
| AF124600 | aroC; aroK; aroB; pepQ | Chorismatsynthase; Shikimatkinae; 3-Dehydrochinatsynthase; mutmaßliche Cytoplasmapeptidase | |
| AF145897 | inhA | | |
| AF145898 | inhA | | Peter, H. et al. "Corynebacterium glutamicum is equipped with four secondary carriers for compatible solutes: Identification, sequencing, and characterization of the proline/ectoine uptake system, ProP, and the ectoine/proline/glycine betaine carrier, EctP," <i>J. Bacteriol.</i> , 180(22):6005-6012 (1998) |
| AJ001436 | ectP | Transport von Ectoine, Glycin, Betain, Prolin | Wehrmann, A. et al. "Different modes of diamminopimelate synthesis and their role in cell wall integrity: A study with <i>Corynebacterium glutamicum</i> ," <i>J. Bacteriol.</i> , 180(12):3159-3165 (1998) |
| AJ004934 | dapD | Tetrahydrodipicolinatsuccinylase (unvollständig) | |
| AJ007732 | ppc; secG; ami; ocd; soxA | Phosphoenolpyruvatcarboxylase; ?; High affinity-Ammonium-Aufnahmeprotein; mutmaßliche Ornithin-cyclodecarboxylase; Sarcosinoxidase | |
| AJ010319 | ftsY; glnB, glnD; srp; amtP | Beteiligt an Zellteilung; PII protein; uridylyltransferase (Uridylyl-enfernendes Enzym); Signalerkennungspartikel; Low affinity-Ammonium-Aufnahmeprotein | Jakoby, M. et al. "Nitrogen regulation in <i>Corynebacterium glutamicum</i> ; Isolation of genes involved in biochemical characterization of corresponding proteins," <i>FEMS Microbiol.</i> , 173(2):303-310 (1999) |
| AJ132968 | cat | Chloramphenicol-acetyltransferase | |

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|-------------------------|------------|--|--|
| AJ224946 | mqo | L-malate: Chinonoxidoreductase | Molenaar, D. et al. "Biochemical and genetic characterization of the membrane-associated malate dehydrogenase (acceptor) from <i>Corynebacterium glutamicum</i> ," <i>Eur. J. Biochem.</i> , 254(2):395-403 (1998) |
| AJ238250 | ndh | NADH-dehydrogenase | Lichinger, T. et al. "Biochemical and biophysical characterization of the cell wall porin of <i>Corynebacterium glutamicum</i> : The channel is formed by a low molecular mass polypeptide," <i>Biochemis/ry</i> , 37(43):15024-15032 (1998) |
| AJ238703 | porA | Porin | Vertes, A.A. et al. "Isolation and characterization of IS31831, a transposable element from <i>Corynebacterium glutamicum</i> ," <i>Mol. Microbiol.</i> , 11(4):739-746 (1994) |
| D17429 | | Transposables Element IS31831 | |
| D84102 | odhA | 2-Oxoglutaratdehydrogenase | Usuda, Y. et al. "Molecular cloning of the <i>Corynebacterium glutamicum</i> (Brevibacterium lactofermentum AJ12036) odhA gene encoding a novel type of 2-oxoglutarate dehydrogenase," <i>Microbiology</i> , 142:3347-3354 (1996) |
| E01358 | hdh; hk | Homoserindehydrogenase; Homoserin-kinase | Katsumata, R. et al. "Production of L-threonine and L-isoleucine," Patent: JP 1987232392-A 1 10/12/87 |
| E01359 | | Stromaufwärts des Startcodons des Homoserinkinase-Gens | Katsumata, R. et al. "Production of L-threonine and L-isoleucine," Patent: JP 1987232392-A 2 10/12/87 |
| E01375 | | Tryptophan-Operon | |
| E01376 | trpL; trpE | Leader-Peptid; Anthranilatsynthase | Matsui, K. et al. "Tryptophan operon, peptide and protein coded thereby, utilization of tryptophan operon gene expression and production of tryptophan," Patent: JP 1987244382-A 1 10/24/87 |
| E01377 | | Promotor- und Operator-Bereiche des Tryptophan-Operons | Matsui, K. et al. "Tryptophan operon, peptide and protein coded thereby, utilization of tryptophan operon gene expression and production of tryptophan," Patent: JP 1987244382-A 1 10/24/87 |
| E03937 | | Biotinsynthase | Hatakeyama, K. et al. "DNA fragment containing gene capable of coding biotin synthetase and its utilization," Patent: JP 1992278088-A 1 10/02/92 |
| E04040 | | Diaminopelargonsäureaminotransferase | Kohama, K. et al. "Gene coding diaminopelargonic acid aminotransferase and desthiobiotin synthetase and its utilization," Patent: JP 1992330284-A 1 11/18/92 |

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|-------------------------|----------|---|--|
| E04041 | | Desthiobiotinsynthetase | Kohama, K. et al. "Gene coding diaminopelargonic acid aminotransferase and desthiobiotin synthetase and its utilization," Patent: JP 1992330284-A 1 11/18/92 |
| E04307 | | Flavum aspartase | Kurusu, Y. et al. "Gene DNA coding aspartase and utilization thereof," Patent: JP 1993030977-A 1 02/09/93 |
| E04376 | | Isocitratlyase | Katsunata, R. et al. "Gene manifestation controlling DNA," Patent: JP 1993056782-A 3 03/09/93 |
| E04377 | | Isocitratlyase N-terminales Fragment | Katsunata, R. et al. "Gene manifestation controlling DNA," Patent: JP 1993056782-A 3 03/09/93 |
| E04484 | | Prephenatdehydratase | Sotouchi, N. et al. "Production of L-phenylalanine by fermentation," Patent: JP 1993076352-A 2 03/30/93 |
| E05108 | | Aspartokinase | Fugono, N. et al. "Gene DNA coding Aspartokinase and its use," Patent: JP 1993184366-A 1 07/27/93 |
| E05112 | | Dihydro-dipichorinatsynthetase | Hatakeyama, K. et al. "Gene DNA coding dihydrodipicolinic acid synthetase and its use," Patent: JP 1993184371-A 1 07/27/93 |
| E05776 | | Diaminopimelinsäuredehydrogenase | Kobayashi, M. et al. "Gene DNA coding Diaminopimelic acid dehydrogenase and its use," Patent: JP 1993284970-A 1 11/02/93 |
| E05779 | | Threoninsynthase | Kohama, K. et al. "Gene DNA coding threonine synthase and its use," Patent: JP 1993284972-A 1 11/02/93 |
| E06110 | | Prephenatdehydratase | Kikuchi, T. et al. "Production of L-phenylalanine by fermentation method," Patent: JP 1993344881-A 1 12/27/93 |
| E06111 | | mutierte Prephenatdehydratase | Kikuchi, T. et al. "Production of L-phenylalanine by fermentation method," Patent: JP 1993344881-A 1 12/27/93 |
| E06146 | | Acetohydroxysäuresynthetase | Inui, M. et al. "Gene capable of coding Acetohydroxy acid synthetase and its use," Patent: JP 1993344893-A 1 12/27/93 |
| E06825 | | Aspartokinase | Suginoto, M. et al. "Mutant aspartokinase gene," patent: JP 1994062866-A 1 03/08/94 |
| E06826 | | mutierte Aspartokinase alpha-Untereinheit | Suginoto, M. et al. "Mutant aspartokinase gene," patent: JP 1994062866-A 1 03/08/94 |

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|---|----------|--|---|
| E06827 | | mutierte Aspartokinase alpha-Untereinheit | Suginoto, M. et al. "Mutant aspartokinase gene," patent: JP 1994062866-A 1 03/08/94 |
| E07701 | secY | | Honno, N. et al. "Gene DNA participating in integration of membranous protein to membrane," Patent: JP 1994169780-A 1 06/21/94 |
| E08177 | | Aspartokinase | Sato, Y. et al. "Genetic DNA capable of coding Aspartokinase released from feedback inhibition and its utilization," Patent: JP 1994261766-A 1 09/20/94 |
| E08178, E08179, E08180, E08181, E08182 | | Durch Rückkopplungshemmung freigesetzte Aspartokinase | Sato, Y. et al. "Genetic DNA capable of coding Aspartokinase released from feedback inhibition and its utilization," Patent: JP 1994261766-A 1 09/20/94 |
| E08232 | | Acetohydroxysäureisomerase | Inui, M. et al. "Gene DNA coding acetohydroxy acid isomerase," Patent: JP 1994277067-A 1 10/04/94 |
| E08234 | secE | | Asai, Y. et al. "Gene DNA coding for translocation machinery of protein," Patent: JP 1994277073-A 1 10/04/94 |
| E08643 | | FT-Aminotransferase und Desthiobiotin-synthetase-Promotorbereich | Hatakeyama, K. et al. "DNA fragment having promoter function in corynebacterium," Patent: JP 1995031476-A 1 02/03/95 |
| E08646 | | Biotinsynthetase | Hatakeyama, K. et al. "DNA fragment having promoter function in corynebacterium," Patent: JP 1995031476-A 1 02/03/95 |
| E08649 | | Aspartase | Kohama, K. et al. "DNA fragment having promoter function in corynebacterium," Patent: JP 1995031478-A 1 02/03/95 |
| E08900 | | Dihydrodipicolinatoreductase | Madori, M. et al. "DNA fragment containing gene coding Dihydrodipicolinate acid reductase and utilization thereof," Patent: JP 1995075578-A 1 03/20/95 |
| E08901 | | Diaminopimelinsäuredecarboxylase | Madori, M. et al. "DNA fragment containing gene coding Diaminopimelate acid decarboxylase and utilization thereof," Patent: JP 1995075579-A 1 03/20/95 |
| E12594 | | Serinhydroxymethyltransferase | Hatakeyama, K. et al. "Production of L-tryptophan," Patent: JP 1997028391-A 1 02/04/97 |
| E12760, E12759, E12758 | | Transposase | Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97 |

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|-------------------------|------------------|--|---|
| E12764 | | Arginyl-tRNA synthetase; Diaminopimelin-säuredecarboxylase | Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97 |
| E12767 | | Dihydrodipicolinsäuresynthetase | Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97 |
| E12770 | | Aspartokinase | Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97 |
| E12773 | | Dihydrodipicolinsäurereductase | Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97 |
| E13655 | | Glucose-6-phosphatedehydrogenase | Hatakeyama, K. et al. "Glucose-6-phosphate dehydrogenase and DNA capable of coding the same," Patent: JP 1997224661-A 1 09/02/97 |
| L01508 | IlvA | Threonindehydratase | Moeckel, B. et al. "Functional and structural analysis of the threonine dehydratase of Corynebacterium glutamicum," <i>J. Bacteriol.</i> , 174:8065-8072 (1992) |
| L07603 | EC 4.2.1.15 | 3-Desoxy-D-arabinoheptulosonat-7-phosphatsynthase | Chen, C. et al. "The cloning and nucleotide sequence of Corynebacterium glutamicum 3-deoxy-D-arabinoheptulosonat-7-phosphate synthase gene," <i>FEMS Microbiol. Lett.</i> , 107:223-230 (1993) |
| L09232 | IlvB; ilvN; ilvC | Acetohydroxysäuresynthase, große Untereinheit; Acetohydroxysäuresynthase kleine Untereinheit; Acetohydroxysäure-isomeroreductase | Keilhauer, C. et al. "Isoleucine synthesis in Corynebacterium glutamicum: molecular analysis of the ilvB-ilvN-ilvC operon," <i>J. Bacteriol.</i> , 175(17):5595-5603 (1993) |
| L18874 | PisM | Phosphoenolpyruvat-Zuckerphosphotransferase | Fouet, A. et al. "Bacillus subtilis sucrose-specific enzyme II of the phosphotransferase system: expression in Escherichia coli and homology to enzymes II from enteric bacteria," <i>PNAS USA</i> , 84(24):8773-8777 (1987); Lee, J.K. et al. "Nucleotide sequence of the gene encoding the Corynebacterium glutamicum mannose enzyme II and analyses of the deduced protein sequence," <i>FEMS Microbiol. Lett.</i> , 119(1-2):137-145 (1994) |
| L27123 | aceB | Malatsynthase | Lee, H-S; et al. "Molecular characterization of aceB, a gene encoding malate synthase in Corynebacterium glutamicum," <i>J. Microbiol. Biotechnol.</i> , 4(4):256-263 (1994) |

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|-------------------------|----------|--------------------------------|--|
| L27126 | | Pyruvatkinase | Jetten, M. S. et al. "Structural and functional analysis of pyruvate kinase from <i>Corynebacterium glutamicum</i> ," <i>Appl. Environ. Microbiol.</i> , 60(7):2501-2507 (1994) |
| L28760 | aceA | Isocitratlyase | Oguiza, J.A. et al. "Molecular cloning, DNA sequence analysis, and characterization of the <i>Corynebacterium diphtheriae</i> dtxR from <i>Brevibacterium lactofermentum</i> ," <i>J. Bacteriol.</i> , 177(2):465-467 (1995) |
| L35906 | dtxr | Diphtherietoxinrepressor | Follettie, M.T. et al. "Molecular cloning and nucleotide sequence of the <i>Corynebacterium glutamicum</i> pheA gene," <i>J. Bacteriol.</i> , 167:695-702 (1986) |
| M13774 | | Prephenaldehydratase | Park, Y.-H. et al. "Phylogenetic analysis of the coryniform bacteria by 56 rRNA sequences," <i>J. Bacteriol.</i> , 169:1801-1806 (1987) |
| M16175 | 5S rRNA | | Sano, K. et al. "Structure and function of the trp operon control regions of <i>Brevibacterium lactofermentum</i> , a glutamic-acid-producing bacterium," <i>Gene</i> , 52:191-200 (1987) |
| M16663 | trpE | Anthranilatsynthase, 5'-Ende | Sano, K. et al. "Structure and function of the trp operon control regions of <i>Brevibacterium lactofermentum</i> , a glutamic-acid-producing bacterium," <i>Gene</i> , 52:191-200 (1987) |
| M16664 | trpA | Tryptophansynthase, 3'-Ende | O'Regan, M. et al. "Cloning and nucleotide sequence of the Phosphoenolpyruvate carboxylase-coding gene of <i>Corynebacterium glutamicum</i> ATCC13032," <i>Gene</i> , 77(2):237-251 (1989) |
| M25819 | | Phosphoenolpyruvatcarboxylase | Roller, C. et al. "Grain-positive bacteria with a high DNA G+C content are characterized by a common insertion within their 23S rRNA genes," <i>J. Gen. Microbiol.</i> , 138:1167-1175 (1992) |
| M85106 | | 23S rRNA-Gen-Insertionssequenz | Roller, C. et al. "Grain-positive bacteria with a high DNA G+C content are characterized by a common insertion within their 23S rRNA genes," <i>J. Gen. Microbiol.</i> , 138:1167-1175 (1992) |
| M85107, M85108 | | 23S rRNA-Gen-Insertionssequenz | |

| GenBank TM Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|--------------------------------------|-------------------------|---|---|
| M89931 | accD; brnQ; yhbW | Beta C-S lyase; Verzweigketten-Aminosäure-Aufnahme-Carrier; hypothetisches Protein yhbW | Rosso, I. et al. "The Corynebacterium glutamicum accD gene encodes a C-S lyase with alpha, beta-elimination activity that degrades aminothylcysteine," <i>J. Bacteriol.</i> , 174(9):2968-2977 (1992); Tauch, A. et al. "Isoleucine uptake in Corynebacterium glutamicum ATCC 13032 is directed by the brnQ gene product," <i>Arch. Microbiol.</i> , 169(4):303-312 (1998) |
| S59299 | tp | Leader-Gen (Promotor) | Herry, D.M. et al. "Cloning of the trp gene cluster from a tryptophan-hyper-producing strain of Corynebacterium glutamicum: identification of a mutation in the trp leader sequence," <i>Appl. Environ. Microbiol.</i> , 59(3):791-799 (1993) |
| U11545 | tpD | Anthranilphosphoribosyltransferase | O'Gara, J.P. and Dunican, L.K. (1994) Complete nucleotide sequence of the Corynebacterium glutamicum ATCC 21850 tpD gene." Thesis, Microbiology Department, University College Galway, Ireland. |
| U13922 | cglIM; cglIR; cglHIR | mutmaßliche Typ II 5-Cytosin-methyltransferase; mutmaßliche Type II Restriktionsendonuklease; mutmaßliche Typ I- oder Typ III Restriktions-endonuklease | Schafer, A. et al. "Cloning and characterization of a DNA region encoding a stress-sensitive restriction system from Corynebacterium glutamicum ATCC 13032 and analysis of its role in intergeneric conjugation with Escherichia coli," <i>J. Bacteriol.</i> , 176(23):7309-7319 (1994); Schafer, A. et al. "The Corynebacterium glutamicum cglIM gene encoding a 5-cytosine in an McrBC-deficient Escherichia coli strain," <i>Gene</i> , 203(2):95-101 (1997) |
| U14965 | recA | | |
| U31224 | ppx | | Ankri, S. et al. "Mutations in the Corynebacterium glutamicum proline biosynthetic pathway: A natural bypass of the proA step," <i>J. Bacteriol.</i> , 178(15):4412-4419 (1996) |
| U31225 | proC | L-Prolin: NADP+ 5-Oxidoreduktase | Ankri, S. et al. "Mutations in the Corynebacterium glutamicum proline biosynthetic pathway: A natural bypass of the proA step," <i>J. Bacteriol.</i> , 178(15):4412-4419 (1996) |
| U31230 | obg; proB; unkdh | ?; Gamma glutamylkinase; ähnlich den D-isomerspezifischen 2-Hydroxysäure-dehydrogenasen | Ankri, S. et al. "Mutations in the Corynebacterium glutamicum proline biosynthetic pathway: A natural bypass of the proA step," <i>J. Bacteriol.</i> , 178(15):4412-4419 (1996) |
| U31281 | bioB | Biotinsynthase | Serebriiskii, I.G., "Two new members of the bio B superfamily: Cloning, sequencing and expression of bio B genes of Methylobacillus flagellatum and Corynebacterium glutamicum," <i>Gene</i> , 175:15-22 (1996) |

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|-------------------------|--|---|--|
| U35023 | thtR; accBC | Thiosulfatschwefeltransferase; Acyl CoA-Carboxylase | Jager, W. et al. "A Corynebacterium glutamicum gene encoding a two-domain protein similar to biotin carboxylases and biotin-carboxyl-carrier proteins," <i>Arch. Microbiol.</i> , 166(2):76-82 (1996) |
| U43535 | cmr | Multidrug-Resistenzprotein | Jager, W. et al. "A Corynebacterium glutamicum gene conferring multidrug resistance in the heterologous host <i>Escherichia coli</i> ," <i>J. Bacteriol.</i> , 179(7):2449-2451 (1997) |
| U43536 | clpB | Hitzeschock-ATP-Bindungsprotein | |
| U53587 | aphA-3 | 3'5"-Aminoglycosidphosphotransferase | |
| U89648 | | Nicht identifizierte Corynebacterium glutamicum-Sequenz, die an der Histidin-biosynthese beteiligt ist, partielle Sequenz | |
| X04960 | trpA; trpB; trpC; trpD; trpE; trpG; trpL | Tryptophanoperon | Matsui, K. et al. "Complete nucleotide and deduced amino acid sequences of the <i>Brevibacterium lactofermentum</i> tryptophan operon," <i>Nucleic Acids Res.</i> , 14(24):10113-10114 (1986) |
| X07563 | lys A | DAP-Decarboxylase (meso-diaminopimelatidecarboxylase, EC 4.1.1.20) | Yeh, P. et al. "Nucleic sequence of the lysA gene of <i>Corynebacterium glutamicum</i> and possible mechanisms for modulation of its expression," <i>Mol. Gen. Genet.</i> , 212(1):112-119 (1988) |
| X14234 | EC 4.1.1.31 | Phosphoenolpyruvatcarboxylase | Eikmanns, B.J. et al. "The Phosphoenolpyruvate carboxylase gene of <i>Corynebacterium glutamicum</i> : Molecular cloning, nucleotide sequence, and expression," <i>Mol. Gen. Genet.</i> , 218(2):330-339 (1989); Lepiniec, L. et al. "Sorghum Phosphoenolpyruvate carboxylase gene family: structure, function and molecular evolution," <i>Plant. Mol. Biol.</i> , 21(3):487-502 (1993) |
| X17313 | fda | Fructose-bisphosphataldolase | Von der Osten, C.H. et al. "Molecular cloning, nucleotide sequence and fine-structural analysis of the <i>Corynebacterium glutamicum</i> fda gene: structural comparison of <i>C. glutamicum</i> fructose-1, 6-bisphosphate aldolase to class I and class II aldolases," <i>Mol. Microbiol.</i> |
| X53993 | dapA | L-2,3-Dihydrodipicolinatsynthetase (EC 4.2.1.52) | Bonnassie, S. et al. "Nucleic sequence of the dapA gene from <i>Corynebacterium glutamicum</i> ," <i>Nucleic Acids Res.</i> , 18(21):6421 (1990) |

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|-------------------------|-------------------------------|---|---|
| X54223 | | AttB-verwandte Stelle | Cianciotto, N. et al. "DNA sequence homology between att B-related sites of Corynebacterium diphtheriae, Corynebacterium ulcerans, Corynebacterium glutamicum, and the attP site of lambdacorynephage," <i>FEMS. Microbiol. Lett.</i> , 66:299-302 (1990) |
| X54740 | argS; lysA | Arginyl-tRNA-synthetase; Diaminopimelat-decarboxylase | Marcel, T. et al. "Nucleotide sequence and organization of the upstream region of the Corynebacterium glutamicum lysA gene," <i>Mol. Microbiol.</i> , 4(11):1819-1830 (1990) |
| X55994 | trpL; trpE | mutmaßliches Leader-Peptid; Anthranilat-synthase-Komponente 1 | Heery, D.M. et al. "Nucleotide sequence of the Corynebacterium glutamicum trpE gene," <i>Nucleic Acids Res.</i> , 18(23):7138 (1990) |
| X56037 | thrC | Threoninsynthase | Han, K.S. et al. "The molecular structure of the Corynebacterium glutamicum threonine synthase gene," <i>Mol. Microbiol.</i> , 4(10):1693-1702 (1990) |
| X56075 | attB-verwandte Stelle | Bindungsstelle | Cianciotto, N. et al. "DNA sequence homology between att B-related sites of Corynebacterium diphtheriae, Corynebacterium ulcerans, Corynebacterium glutamicum, and the attP site of lambdacorynephage," <i>FEMS. Microbiol. Lett.</i> , 66:299-302 (1990) |
| X57226 | lysC-alpha; lysC-beta; asd | Aspartokinase-alpha-Untereinheit; Aspartokinase-beta-Untereinheit; Aspartat-beta-semialdehyddehydrogenase | Kalinowski, J. et al. "Genetic and biochemical analysis of the Aspartokinase from Corynebacterium glutamicum," <i>Mol. Microbiol.</i> , 5(5):1197-1204 (1991); Kalinowski, J. et al. "Aspartokinase genes lysC alpha and lysC beta overlap and are adjacent to the aspartate beta-semialdehyde dehydrogenase gene asd in Corynebacterium glutamicum," <i>Mol. Gen. Genet.</i> , 224(3):317-324 (1990) |
| X59403 | gap;pgk; tpi | Glyceraldehyde-3-phosphat; Phosphoglyceratkinase; Triosephosphat-isomerase | Eikmanns, B.J. "Identification, sequence analysis, and expression of a Corynebacterium glutamicum gene cluster encoding the three glycolytic enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate kinase, and triosephosphate isomerase," <i>J. Bacteriol.</i> , 174(19):6076-6086 (1992) |
| X59404 | gdh | Glutamatdehydrogenase | Bornmann, E.R. et al. "Molecular analysis of the Corynebacterium glutamicum gdh gene encoding glutamate dehydrogenase," <i>Mol. Microbiol.</i> , 6(3):317-326 (1992) |
| X60312 | lysI | L-Lysinpermease | Seep-Feldhaus, A.H. et al. "Molecular analysis of the Corynebacterium glutamicum lysI gene involved in lysine uptake," <i>Mol. Microbiol.</i> , 5(12):2995-3005 (1991) |

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|-------------------------|----------|-----------------------------------|--|
| X66078 | cop1 | Ps1 protein | Joliff, G. et al. "Cloning and nucleotide sequence of the csp1 gene encoding PS1, one of the two major secreted proteins of <i>Corynebacterium glutamicum</i> : The deduced N-terminal region of PS1 is similar to the <i>Mycobacterium</i> antigen 85 complex," <i>Mol. Microbiol.</i> , 6(16):2349-2362 (1992) |
| X66112 | glt | Citrat synthase | Eikmanns, B.J. et al. "Cloning sequence, expression and transcriptional analysis of the <i>Corynebacterium glutamicum</i> gltA gene encoding citrate synthase," <i>Microbiol.</i> , 140:1817-1828 (1994) |
| X67737 | dapB | Dihydrodipicolinate reductase | Peyret, J.L. et al. "Characterization of the cspB gene encoding PS2, an ordered surface-layer protein in <i>Corynebacterium glutamicum</i> ," <i>Mol. Microbiol.</i> , 9(1):97-109 (1993) |
| X69103 | csp2 | Oberflächenprotein PS2 | Bonamy, C. et al. "Identification of IS1206, a <i>Corynebacterium glutamicum</i> IS3-related insertion sequence and phylogenetic analysis," <i>Mol. Microbiol.</i> , 14(3):571-581 (1994) |
| X69104 | | IS3-verwandtes Insertions-element | Patek, M. et al. "Leucine synthesis in <i>Corynebacterium glutamicum</i> : enzyme activities, structure of leuA, and effect of leuA inactivation on lysine synthesis," <i>Appl. Environ. Microbiol.</i> , 60(1):133-140 (1994) |
| X70959 | leuA | Isopropylmalat synthase | Eikmanns, B.J. et al. "Cloning sequence analysis, expression, and inactivation of the <i>Corynebacterium glutamicum</i> icd gene encoding isocitrate dehydrogenase and biochemical characterization of the enzyme," <i>J. Bacteriol.</i> , 177(3):774-782 (1995) |
| X71489 | icd | Isocitratdehydrogenase (NADP+) | |
| X72855 | GDHA | Glutamatdehydrogenase (NADP+) | Heery, D.M. et al. "A sequence from a tryptophan-hyperproducing strain of <i>Corynebacterium glutamicum</i> encoding resistance to 5-methyltryptophan," <i>Biochem. Biophys. Res. Commun.</i> , 201(3):1255-1262 (1994) |
| X75083, X70584 | mttA | 5-Methyltryptophanresistenz | Fitzpatrick, R. et al. "Construction and characterization of recA mutant strains of <i>Corynebacterium glutamicum</i> and <i>Brevibacterium lactofermentum</i> ," <i>Appl. Microbiol. Biotechnol.</i> , 42(4):575-580 (1994) |
| X75085 | recA | | |

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|-------------------------|---------------------------|-------------------------------------|--|
| X75504 | aceA; thiX | partielle Isocitratlyase; ? | Reinscheid, D.J. et al. "Characterization of the isocitrate lyase gene from <i>Corynebacterium glutamicum</i> and biochemical analysis of the enzyme," <i>J. Bacteriol.</i> , 176(12):3474-3483 (1994) |
| X76875 | | ATPase beta-Untereinheit | Ludwig, W. et al. "Phylogenetic relationships of bacteria based on comparative sequence analysis of elongation factor Tu and ATP-synthase beta-subunit genes," <i>Antonie Van Leeuwenhoek</i> , 64:285-305 (1993) |
| X77034 | tuf | Elongationsfaktor Tu | Ludwig, W. et al. "Phylogenetic relationships of bacteria based on comparative sequence analysis of elongation factor Tu and ATP-synthase beta-subunit genes," <i>Antonie Van Leeuwenhoek</i> , 64:285-305 (1993) |
| X77384 | recA | | Bilman-Jacobe, H. "Nucleotide sequence of a recA gene from <i>Corynebacterium glutamicum</i> ," <i>DNA Seq.</i> , 4(6):403-404 (1994) |
| X78491 | aceB | Malatsynthase | Reinscheid, D.J. et al. "Malate synthase from <i>Corynebacterium glutamicum</i> pta-ack operon encoding phosphotransacetylase: sequence analysis," <i>Microbiology</i> , 140:3099-3108 (1994) |
| X80629 | 16S rDNA | 16S ribosomale RNA | Rainey, F.A. et al. "Phylogenetic analysis of the genera <i>Rhodococcus</i> and <i>Norcardia</i> and evidence for the evolutionary origin of the genus <i>Norcardia</i> from within the radiation of <i>Rhodococcus</i> species," <i>Microbiol.</i> , 141:523-528 (1995) |
| X81191 | gluA; gluB; gluC; gluD | Glutamat-Aufnahmesystem | Kronmeyer, W. et al. "Structure of the gluABCD cluster encoding the glutamate uptake system of <i>Corynebacterium glutamicum</i> ," <i>J. Bacteriol.</i> , 177(5):1152-1158 (1995) |
| X81379 | dapE | Succinyldiaminopimelatesuccinylase | Wehrmann, A. et al. "Analysis of different DNA fragments of <i>Corynebacterium glutamicum</i> complementing dapE of <i>Escherichia coli</i> ," <i>Microbiology</i> , 40:3349-56 (1994) |
| X82061 | 16S rDNA | 16S ribosomale RNA | Ruiny, R. et al. "Phylogeny of the genus <i>Corynebacterium</i> deduced from analyses of small-subunit ribosomal DNA sequences," <i>Int. J. Syst. Bacteriol.</i> , 45(4):740-746 (1995) |
| X82928 | asd; lysC | Aspartatsemialdehyddehydrogenase; ? | Serebrijski, I. et al. "Multicopy suppression by asd gene and osmotic stress-dependent complementation by heterologous proA in proA mutants," <i>J. Bacteriol.</i> , 177(24):7255-7260 (1995) |

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|-------------------------|---------------------------------|--|---|
| X82929 | proA | Gamma-glutamylphosphatreduktase | Serebrjski, I. et al. "Multicopy suppression by asd gene and osmotic stress-dependent complementation by heterologous proA in proA mutants," <i>J. Bacteriol.</i> , 177(24):7255-7260 (1995) |
| X84257 | 16S rDNA | 16S ribosomale RNA | Pascual, C. et al. "Phylogenetic analysis of the genus <i>Corynebacterium</i> based on 16S rRNA gene sequences," <i>Int. J. Syst. Bacteriol.</i> , 45(4):724-728 (1995) |
| X85965 | aroP, dapE | aromatische Aminosäurepermease; ? | Wehrmann, A. et al. "Functional analysis of sequences adjacent to dapE of <i>Corynebacterium glutamicum</i> proline reveals the presence of aroP, which encodes the aromatic amino acid transporter," <i>J. Bacteriol.</i> , 177(20):5991-5993 (1995) |
| X86157 | argB; argC; argD; argF; argJ | Acetylglutaminkinase; N-acetyl-gamma-glutamyl-phosphatreduktase; Acetylornithinaminotransferase; Ornithin-carbamoyltransferase; Glutamat-N-acetyltransferase | Sakanyan, V. et al. "Genes and enzymes of the acetyl cycle of arginine biosynthesis in <i>Corynebacterium glutamicum</i> : enzyme evolution in the early steps of the arginine pathway," <i>Microbiology</i> , 142:99-108 (1996) |
| X89084 | pta; ackA | Phosphatase/transferase; Acetatekinase | Reinscheid, D.J. et al. "Cloning, sequence analysis, expression and inactivation of the <i>Corynebacterium glutamicum</i> pta-ack operon encoding phosphotransacetylase and acetate kinase," <i>Microbiology</i> , 145:503-513 (1999) |
| X89850 | attB | Bindungsstelle | Le Marrec, C. et al. "Genetic characterization of site-specific integration functions of phi AAU2 infecting 'Arthrobacter aureus C70,'" <i>J. Bacteriol.</i> , 178(7):1996-2004 (1996) |
| X90356 | | Promotorfragment F1 | Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996) |
| X90357 | | Promotorfragment F2 | Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996) |
| X90358 | | Promotorfragment F10 | Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996) |

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|-------------------------|----------|------------------------|--|
| X90359 | | Promotorfragment F13 | Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996) |
| X90360 | | Promotorfragment F22 | Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996) |
| X90361 | | Promotorfragment F34 | Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996) |
| X90362 | | Promotorfragment F37 | Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996) |
| X90363 | | Promotorfragment F45 | Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996) |
| X90364 | | Promotorfragment F64 | Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996) |
| X90365 | | Promotorfragment F75 | Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996) |
| X90366 | | Promotorfragment PF101 | Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996) |
| X90367 | | Promotorfragment PF104 | Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996) |
| X90368 | | Promotorfragment PF109 | Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996) |

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|-------------------------|---------------------|---|--|
| X93513 | amt | Ammonium-Transportsystem | Siewe, R.M. et al. "Functional and genetic characterization of the (methyl) ammonium uptake carrier of <i>Corynebacterium glutamicum</i> ," <i>J. Biol. Chem.</i> 271(10):5398-5403 (1996) |
| X93514 | betP | Glycin-Betain-Transportsystem | Peter, H. et al. "Isolation, characterization, and expression of the <i>Corynebacterium glutamicum</i> betP gene, encoding the transport system for the compatible solute glycine betaine," <i>J. Bacteriol.</i> 178(17):5229-5234 (1996) |
| X95649 | orf4 | | Patek, M. et al. "Identification and transcriptional analysis of the dapB-ORF2-dapA-ORF4 operon of <i>Corynebacterium glutamicum</i> , encoding two enzymes involved in L-lysine synthesis," <i>Biotechnol. Lett.</i> 19:1113-1117 (1997) |
| X96471 | lysE; lysG | Lysinexportprotein; Lysinexportregulator-protein | Vrijic, M. et al. "A new type of transporter with a new type of cellular function: L-lysine export from <i>Corynebacterium glutamicum</i> ," <i>Mol. Microbiol.</i> 22(5):815-826 (1996) |
| X96580 | panB; panC; xylB | 3-Methyl-2-oxobutanoatehydroxymethyl-transferase; Pantoat-beta-alaninligase; Xylulokinase | Sahn, H. et al. "D-pantothenate synthesis in <i>Corynebacterium glutamicum</i> and use of panBC and genes encoding L-valine synthesis for D-pantothenate overproduction," <i>Appl. Environ. Microbiol.</i> 65(5):1973-1979 (1999) |
| X96962 | | Insertionssequenz IS1207 und Transposase | |
| X99289 | | Elongationsfaktor P | Ramos, A. et al. "Cloning, sequencing and expression of the gene encoding elongation factor P in the amino-acid producer <i>Brevibacterium lactofermentum</i> (<i>Corynebacterium glutamicum</i> ATCC 13869)," <i>Gene</i> 198:217-222 (1997) |
| Y00140 | thrB | Homoserinkinase | Mateos, L.M. et al. "Nucleotide sequence of the homoserine kinase (thrB) gene of the <i>Brevibacterium lactofermentum</i> ," <i>Nucleic Acids Res.</i> 15(9):3922 (1987) |
| Y00151 | ddh | Meso-diaminopimelat-D-dehydrogenase (EC 1.4.1.16) | Ishino, S. et al. "Nucleotide sequence of the meso-diaminopimelate D-dehydrogenase gene from <i>Corynebacterium glutamicum</i> ," <i>Nucleic Acids Res.</i> 15(9):3917 (1987) |
| Y00476 | thrA | Homoserindehydrogenase | Mateos, L.M. et al. "Nucleotide sequence of the homoserine dehydrogenase (thrA) gene of the <i>Brevibacterium lactofermentum</i> ," <i>Nucleic Acids Res.</i> 15(24):10598 (1987) |

| GenBank TM Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|--------------------------------------|--------------------------|---|--|
| Y00546 | hom; thrB | Homoserindehydrogenase; Homoserin-kinase | Peoples, O.P. et al. "Nucleotide sequence and fine structural analysis of the Corynebacterium glutamicum hom-thrB operon," <i>Mol. Microbiol.</i> , 2(1):63-72 (1988) |
| Y08964 | murC; ftsQ/divD; ftsZ | UPD-N-Acetylmuramatalanilgase; Teilungsinitiationsprotein oder Zellteilungsprotein; Zellteilungsprotein | Honrubia, M.P. et al. "Identification, characterization, and chromosomal organization of the ftsZ gene from Brevibacterium lactofermentum," <i>Mol. Gen. Genet.</i> , 259(1):97-104 (1998) |
| Y09163 | putP | High affinity-Prolintransportsystem | Peter, H. et al. "Isolation of the putP gene of Corynebacterium glutamicum-proline and characterization of a low-affinity uptake system for compatible solutes," <i>Arch. Microbiol.</i> , 168(2):143-151 (1997) |
| Y09548 | pyc | Pyruvatcarboxylase | Peters-Wendisch, P.G. et al. "Pyruvate carboxylase from Corynebacterium glutamicum: characterization, expression and inactivation of the pyc gene," <i>Microbiology</i> , 144:915-927 (1998) |
| Y09578 | leuB | 3-Isopropylmalaldehydhydrogenase | Patek, M. et al. "Analysis of the leuB gene from Corynebacterium glutamicum," <i>Appl. Microbiol. Biotechnol.</i> , 50(1):42-47 (1998) |
| Y12472 | | Bindungsstelle Bacteriophage Phi-16 | Moreau, S. et al. "Site-specific integration of corynephage Phi-16: The construction of an integration vector," <i>Microbiol.</i> , 145:539-548 (1999) |
| Y12537 | proP | Prolin/Ectoin-Aufnahmesystemprotein | Peter, H. et al. "Corynebacterium glutamicum is equipped with four secondary carriers for compatible solutes: Identification, sequencing, and characterization of the proline/ectoine uptake system, ProP, and the ectoine/proline/glycine betaine carrier, EctP," <i>J. Bacteriol.</i> , 180(22):6005-6012 (1998) |
| Y13221 | glnA | Glutaminsynthetase I | Jakoby, M. et al. "Isolation of Corynebacterium glutamicum glnA gene encoding glutamine synthetase I," <i>FEMS Microbiol. Lett.</i> , 154(1):81-88 (1997) |
| Y16642 | lpd | Dihydrolipoamiddehydrogenase | |
| Y18059 | | Bindungsstelle Corynephage 304L | Moreau, S. et al. "Analysis of the integration functions of φ304L: An integrase module among corynephages," <i>Virology</i> , 255(1):150-159 (1999) |
| Z21501 | argS; lysA | Arginyl-tRNA-Synthetase; Diaminopimelatdecarboxylase (partiell) | Oguiza, J.A. et al. "A gene encoding arginyl-tRNA synthetase is located in the upstream region of the lysA gene in Brevibacterium lactofermentum: Regulation of argS-lysA cluster expression by arginine," <i>J. Bacteriol.</i> , 175(22):7356-7362 (1993) |

| GenBank™ Zugangs-Nr. | Gen-Name | Genfunktion | Literaturstellen |
|-------------------------|------------|---|--|
| Z21502 | dapA; dapB | Dihydrodipicolinatsynthase; Dihydrodipicolinatreduktase | Pisabarro, A. et al. "A cluster of three genes (dapA, orf2, and dapB) of Brevibacterium lactofermentum encodes dihydrodipicolinate reductase, and a third polypeptide of unknown function," <i>J. Bacteriol.</i> , 175(9):2743-2749 (1993) |
| Z29563 | thrC | Threoninsynthase | Malumbres, M. et al. "Analysis and expression of the thrC gene of the encoded threonine synthase," <i>Appl. Environ. Microbiol.</i> , 60(7):2209-2219 (1994) |
| Z46753 | 16S rDNA | Gene für 16S ribosomale RNA | |
| Z49822 | sigA | SigA-Signalfaktor | Oguiza, J.A. et al. "Multiple sigma factor genes in Brevibacterium lactofermentum: Characterization of sigA and sigB," <i>J. Bacteriol.</i> , 178(2):550-553 (1996) |
| Z49823 | galE; dtxR | Katalytische Aktivität UDP-Galactose 4-epimerase; Diphtherietoxin-regulatorisches Protein | Oguiza, J.A. et al. "The galE gene encoding the UDP-galactose 4-epimerase of Brevibacterium lactofermentum is coupled transcriptionally to the dndR gene," <i>Gene</i> , 177:103-107 (1996) |
| Z49824 | orf1; sigB | ?; SigB-Signalfaktor | Oguiza, J.A. et al. "Multiple sigma factor genes in Brevibacterium lactofermentum: Characterization of sigA and sigB," <i>J. Bacteriol.</i> , 178(2):550-553 (1996) |
| Z66534 | | Transposase | Correia, A. et al. "Cloning and characterization of an IS-like element present in the genome of Brevibacterium lactofermentum ATCC 13869," <i>Gene</i> , 170(1):91-94 (1996) |

- 1) Eine Sequenz für dieses Gen wurde in den angegebenen Literaturstellen veröffentlicht. Die von den Erfindern der vorliegenden Erfindung erhaltene Sequenz ist jedoch erheblich länger als die veröffentlichte Version. Man nimmt an, daß die veröffentlichte Version auf einem inkorrekten Startcodon beruht und somit nur ein Fragment des tatsächlichen codierenden Bereichs darstellt.

TABELLE 3: Corynebacterium- und Brevibacterium-Stämme, die sich in der erfindungsgemäßen Praxis einsetzen lassen.

| Gattung | Art | ATCC | FERM | NRRL | CECT | NCIMB | CBS | NCTC | DSMZ |
|----------------|-------------|-------|------|--------|------|-------|-----|------|------|
| Brevibacterium | ammoniaenes | 21054 | | | | | | | |
| Brevibacterium | ammoniaenes | 19350 | | | | | | | |
| Brevibacterium | ammoniaenes | 19351 | | | | | | | |
| Brevibacterium | ammoniaenes | 19352 | | | | | | | |
| Brevibacterium | ammoniaenes | 19353 | | | | | | | |
| Brevibacterium | ammoniaenes | 19354 | | | | | | | |
| Brevibacterium | ammoniaenes | 19355 | | | | | | | |
| Brevibacterium | ammoniaenes | 19356 | | | | | | | |
| Brevibacterium | ammoniaenes | 21055 | | | | | | | |
| Brevibacterium | ammoniaenes | 21077 | | | | | | | |
| Brevibacterium | ammoniaenes | 21553 | | | | | | | |
| Brevibacterium | ammoniaenes | 21580 | | | | | | | |
| Brevibacterium | ammoniaenes | 39101 | | | | | | | |
| Brevibacterium | butanicum | 21196 | | | | | | | |
| Brevibacterium | divaricatum | 21792 | P928 | | | | | | |
| Brevibacterium | flavum | 21474 | | | | | | | |
| Brevibacterium | flavum | 21129 | | | | | | | |
| Brevibacterium | flavum | 21518 | | | | | | | |
| Brevibacterium | flavum | | | B11474 | | | | | |
| Brevibacterium | flavum | | | B11472 | | | | | |
| Brevibacterium | flavum | 21127 | | | | | | | |
| Brevibacterium | flavum | 21128 | | | | | | | |
| Brevibacterium | flavum | 21427 | | | | | | | |
| Brevibacterium | flavum | 21475 | | | | | | | |
| Brevibacterium | flavum | 21517 | | | | | | | |
| Brevibacterium | flavum | 21528 | | | | | | | |
| Brevibacterium | flavum | 21529 | | | | | | | |

| Gattung | Art | ATCC | FERM | NRRL | CECT | NCIMB | CBS | NCTC | DSMZ |
|----------------|------------------|-------|------|--------|------|-------|--------|------|------|
| Brevibacterium | flavum | | | B11477 | | | | | |
| Brevibacterium | flavum | | | B11478 | | | | | |
| Brevibacterium | flavum | 21127 | | | | | | | |
| Brevibacterium | flavum | | | B11474 | | | | | |
| Brevibacterium | healii | 15527 | | | | | | | |
| Brevibacterium | ketoglutamicum | 21004 | | | | | | | |
| Brevibacterium | ketoglutamicum | 21089 | | | | | | | |
| Brevibacterium | ketosoreductum | 21914 | | | | | | | |
| Brevibacterium | lactofermentum | | | | 70 | | | | |
| Brevibacterium | lactofermentum | | | | 74 | | | | |
| Brevibacterium | lactofermentum | | | | 77 | | | | |
| Brevibacterium | lactofermentum | 21798 | | | | | | | |
| Brevibacterium | lactofermentum | 21799 | | | | | | | |
| Brevibacterium | lactofermentum | 21800 | | | | | | | |
| Brevibacterium | lactofermentum | 21801 | | | | | | | |
| Brevibacterium | lactofermentum | | | B11470 | | | | | |
| Brevibacterium | lactofermentum | | | B11471 | | | | | |
| Brevibacterium | lactofermentum | 21086 | | | | | | | |
| Brevibacterium | lactofermentum | 21420 | | | | | | | |
| Brevibacterium | lactofermentum | 21086 | | | | | | | |
| Brevibacterium | lactofermentum | 31269 | | | | | | | |
| Brevibacterium | linens | 9174 | | | | | | | |
| Brevibacterium | linens | 19391 | | | | | | | |
| Brevibacterium | linens | 8377 | | | | | | | |
| Brevibacterium | paraffinolyticum | | | | | 11160 | | | |
| Brevibacterium | spec. | | | | | | 717.73 | | |
| Brevibacterium | spec. | | | | | | 717.73 | | |
| Brevibacterium | spec. | 14604 | | | | | | | |
| Brevibacterium | spec. | 21860 | | | | | | | |
| Brevibacterium | spec. | 21864 | | | | | | | |

| Gattung | Art | ATCC | FERM | NRRL | CECT | NCIMB | CBS | NCTC | DSMZ |
|-----------------|------------------|-------|------|--------|------|-------|-----|------|------|
| Brevibacterium | spec. | 21865 | | | | | | | |
| Brevibacterium | spec. | 21866 | | | | | | | |
| Brevibacterium | spec. | 19240 | | | | | | | |
| Corynebacterium | acetoacidophilum | 21476 | | | | | | | |
| Corynebacterium | acetoacidophilum | 13870 | | | | | | | |
| Corynebacterium | acetoglutamicum | | | B11473 | | | | | |
| Corynebacterium | acetoglutamicum | | | B11475 | | | | | |
| Corynebacterium | acetoglutamicum | 15806 | | | | | | | |
| Corynebacterium | acetoglutamicum | 21491 | | | | | | | |
| Corynebacterium | acetoglutamicum | 31270 | | | | | | | |
| Corynebacterium | acetophilum | | | B3671 | | | | | |
| Corynebacterium | annonniagenes | 6872 | | | | | | 2399 | |
| Corynebacterium | annonniagenes | 15511 | | | | | | | |
| Corynebacterium | fujikense | 21496 | | | | | | | |
| Corynebacterium | glutamicum | 14067 | | | | | | | |
| Corynebacterium | glutamicum | 39137 | | | | | | | |
| Corynebacterium | glutamicum | 21254 | | | | | | | |
| Corynebacterium | glutamicum | 21255 | | | | | | | |
| Corynebacterium | glutamicum | 31830 | | | | | | | |
| Corynebacterium | glutamicum | 13032 | | | | | | | |
| Corynebacterium | glutamicum | 14305 | | | | | | | |
| Corynebacterium | glutamicum | 15455 | | | | | | | |
| Corynebacterium | glutamicum | 13058 | | | | | | | |
| Corynebacterium | glutamicum | 13059 | | | | | | | |
| Corynebacterium | glutamicum | 13060 | | | | | | | |
| Corynebacterium | glutamicum | 21492 | | | | | | | |
| Corynebacterium | glutamicum | 21513 | | | | | | | |
| Corynebacterium | glutamicum | 21526 | | | | | | | |
| Corynebacterium | glutamicum | 21543 | | | | | | | |
| Corynebacterium | glutamicum | 13287 | | | | | | | |

| Gattung | Art | ATCC | FERM | NRRL | CECT | NCIMB | CBS | NCTC | DSMZ |
|-----------------|------------|-------|------|------|------|-------|-----|------|------|
| Corynebacterium | glutamicum | 21851 | | | | | | | |
| Corynebacterium | glutamicum | 21253 | | | | | | | |
| Corynebacterium | glutamicum | 21514 | | | | | | | |
| Corynebacterium | glutamicum | 21516 | | | | | | | |
| Corynebacterium | glutamicum | 21299 | | | | | | | |
| Corynebacterium | glutamicum | 21300 | | | | | | | |
| Corynebacterium | glutamicum | 39684 | | | | | | | |
| Corynebacterium | glutamicum | 21488 | | | | | | | |
| Corynebacterium | glutamicum | 21649 | | | | | | | |
| Corynebacterium | glutamicum | 21650 | | | | | | | |
| Corynebacterium | glutamicum | 19223 | | | | | | | |
| Corynebacterium | glutamicum | 13869 | | | | | | | |
| Corynebacterium | glutamicum | 21157 | | | | | | | |
| Corynebacterium | glutamicum | 21158 | | | | | | | |
| Corynebacterium | glutamicum | 21159 | | | | | | | |
| Corynebacterium | glutamicum | 21355 | | | | | | | |
| Corynebacterium | glutamicum | 31808 | | | | | | | |
| Corynebacterium | glutamicum | 21674 | | | | | | | |
| Corynebacterium | glutamicum | 21562 | | | | | | | |
| Corynebacterium | glutamicum | 21563 | | | | | | | |
| Corynebacterium | glutamicum | 21564 | | | | | | | |
| Corynebacterium | glutamicum | 21565 | | | | | | | |
| Corynebacterium | glutamicum | 21566 | | | | | | | |
| Corynebacterium | glutamicum | 21567 | | | | | | | |
| Corynebacterium | glutamicum | 21568 | | | | | | | |
| Corynebacterium | glutamicum | 21569 | | | | | | | |
| Corynebacterium | glutamicum | 21570 | | | | | | | |
| Corynebacterium | glutamicum | 21571 | | | | | | | |
| Corynebacterium | glutamicum | 21572 | | | | | | | |
| Corynebacterium | glutamicum | 21573 | | | | | | | |

| Gattung | Art | ATCC | FERM | NRRL | CECT | NCIMB | CBS | NCTC | DSMZ |
|-----------------|---------------|-------|-------|--------|------|-------|-----|------|------|
| Corynebacterium | glutamicum | 21579 | | | | | | | |
| Corynebacterium | glutamicum | 19049 | | | | | | | |
| Corynebacterium | glutamicum | 19050 | | | | | | | |
| Corynebacterium | glutamicum | 19051 | | | | | | | |
| Corynebacterium | glutamicum | 19052 | | | | | | | |
| Corynebacterium | glutamicum | 19053 | | | | | | | |
| Corynebacterium | glutamicum | 19054 | | | | | | | |
| Corynebacterium | glutamicum | 19055 | | | | | | | |
| Corynebacterium | glutamicum | 19056 | | | | | | | |
| Corynebacterium | glutamicum | 19057 | | | | | | | |
| Corynebacterium | glutamicum | 19058 | | | | | | | |
| Corynebacterium | glutamicum | 19059 | | | | | | | |
| Corynebacterium | glutamicum | 19060 | | | | | | | |
| Corynebacterium | glutamicum | 19185 | | | | | | | |
| Corynebacterium | glutamicum | 13286 | | | | | | | |
| Corynebacterium | glutamicum | 21515 | | | | | | | |
| Corynebacterium | glutamicum | 21527 | | | | | | | |
| Corynebacterium | glutamicum | 21544 | | | | | | | |
| Corynebacterium | glutamicum | 21492 | | | | | | | |
| Corynebacterium | glutamicum | | | B8183 | | | | | |
| Corynebacterium | glutamicum | | | B8182 | | | | | |
| Corynebacterium | glutamicum | | | B12416 | | | | | |
| Corynebacterium | glutamicum | | | B12417 | | | | | |
| Corynebacterium | glutamicum | | | B12418 | | | | | |
| Corynebacterium | glutamicum | | | B11476 | | | | | |
| Corynebacterium | glutamicum | 21608 | | | | | | | |
| Corynebacterium | lilium | | P973 | | | | | | |
| Corynebacterium | nitrilophilus | 21419 | | | | 11594 | | | |
| Corynebacterium | spec. | | P4445 | | | | | | |
| Corynebacterium | spec. | | P4446 | | | | | | |

| Gattung | Art | ATCC | FERM | NRRL | CECT | NCIMB | CBS | NCTC | DSMZ |
|-----------------|-------|-------|------|------|------|-------|-----|------|-------|
| Corynebacterium | spec. | 31088 | | | | | | | |
| Corynebacterium | spec. | 31089 | | | | | | | |
| Corynebacterium | spec. | 31090 | | | | | | | |
| Corynebacterium | spec. | 31090 | | | | | | | |
| Corynebacterium | spec. | 31090 | | | | | | | |
| Corynebacterium | spec. | 15954 | | | | | | | 20145 |
| Corynebacterium | spec. | 21857 | | | | | | | |
| Corynebacterium | spec. | 21862 | | | | | | | |
| Corynebacterium | spec. | 21863 | | | | | | | |

ATCC: American Type Culture Collection, Rockville, MD, USA
 FERM: Fermentation Research Institute, Chiba, Japan
 NRRL: ARS Culture Collection, Northern Regional Research Laboratory, Peoria, IL, USA
 CECT: Coleccion Espanola de Cultivos Tipo, Valencia, Spain
 NCIMB: National Collection of Industrial and Marine Bacteria Ltd., Aberdeen, UK
 CBS: Centraalbureau voor Schimmelcultures, Baarn, NL
 NCTC: National Collection of Type Cultures, London, UK
 DSMZ: Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, Germany

Siehe Sugawara, H. et al. (1993) World directory of collections of cultures of microorganisms: Bacteria, fungi and yeasts (4. Aufl.), World federation for culture collections world data center on microorganisms, Saimata, Japan.

RXN01703 translatiert von RXN01703 (3504) von 1 bis 1113

VAWPSNAKEKLFHWHYWWQAHYLDCLVDAARRRTTKARRDRIRDITIRGISVRNVGKLTSNRYDDKAWLALALGRAGKVRKV
RTPKSLPSLEQNIVDGIDSLTGVLWPWSGETFYNVPSNGPAAIMMARTDRLDEAMKITDWIFDNLIDGDLVMDGLRMRMHGP
ELVRSIHYPYCQGAIGACLEIALKLRRERAGLTITVVDHWSADKAEDSLKYFAHIHAVVQAVSRKMTNFHGVIDWDTGDGDGG
LFKGILVRYLADVAILRPDDSPNTNRETKKIAARLVLESAESVWNHRLEVDGLPVFATDWTTDARLPQNFGLSSSSLSLVSVV
RVDERDLSVQLSGWMLMEAAAKVAEELNNGNSYTGRSR

RXN01703 - 5'-Region

GTTAGACAAATGGGTAAACAGAGCTGACCTAGCGGAATCCGCCATCAACGAAAGGCATTCCGCGAGGGTTTGGGGTCTGCCTC
GAACAAATCTTGGGTTT

RXN01703 - kodierende Region

GTGGCATGGCCATCCAACGCCAAAGAAAACTGTTTATCCACTGGCAAGCGCATTATCTAGACTGCCTGGT
GGATGCTGCTCGTCGACGCACCACAAAGGCCCGTCGCGACCGCATCAGGGACACCATCCGCGGCATTTCCGGTGCCTGCAATGTGG
GCAAGCTGACCTCGAATCGTTATTACGACGACAAAGCTTGGCTGGCCCTTGCTCTTGGGCGTGCCGGAAGAGTGCGAAAGGTG
CGCACACCAAAATCATTGCCCTCGTTGGAACAAACATCGTCGATGGCATTGATTCCCTTACTGGTGTGCTGCCGTGGCGTTC
CGGCGAAACCTTCTACAACGTTCCCTCCAACGGTCTGCTGCGATCATGATGGCCCGCACCGACCGTTTGGACGAGGCTATGA
AAATCACCGATTGGATTTTGGACAACCTGATCGATGGCGACGGCCTTGTGATGGACGGATTGCGCATGCGCATGCACGGACCT
GAGCTTGTCCGTTCCATCCACCCGTATTGCCAAGGTGTGCGCCATTGGTGCCTGTTTGGAAATTGCTCTCAAACCTGCGTGAGCG
GCAGGCTTGACCACTACTGTGGTGGATCACTGGTGGATGCCGATAAGGCAGAAGACTCCCTCAAATACTTTGCACACATCC
ACGCTGTGGTTTCAAGGCTGTGTCGCGGAAGATGACCAACTTCCACGGCGTTATTGATTGGGACACCGGTGACGGCGACGGCGGT
TTGTTCAAGGGCATTTTGGTCCGCTATTTAGCTGATGTGGCCATCCGCTGCGCTGACGATTACCAACCAACCGGGAACCAA
AAAGATTGCAGCACGCTGGTACTGGAATCGGCGGAAAGCGTATGGAACACCGATTGGAAGTTGATGGCCTTCCGGTATTG
CCACAGACTGGACAACGGATGCACGCCTGCCACAAAACCTTTGGTTTGGAGTTCCTCTAGTTTGAGCGATCTGGTGGAGTGTGTG
CGCGTGGATGAACGTGATCTGTCCGTGCAATTGTCCGGTTGGATGCTCATGGAAGCAGCAGCGAAAGTGCCGAAGAAGTGA
AAACAACGGCAATAGTTACACCGGTGCGTCCCGA

RXN01703 - 3'-Region

TAGCCCCGATAGTGTATGTGCTG

RXN01694 translatiert von RXN01694 (8929) von 1 bis 900

MQKITPNIWCQGTADAAEFYVNAFSEFPGGAEVLTTVKYPEAGLLDFQEPFAGKTLTVELAISGFKIILINAGEEFTPNPSI
SFMVNFDAVRDENAHEHLDAVWEKLHEGGSTLMPVDTPFSEYYGWVQDKYGVSWQLMLSRPEEKPGPAVPTLLFGGAAQNO
AGPAQENYVEVPNSQLGDRAPYGQQTGPATPEALMFSQFQLDGQWIFAMDSGVEQDFTFSEGVSLMYEAHQEELDAIWNAL
SAVPEAEACGWLKDKFGVSWQIVPDNMEELMAKPGAYEKLLAMKKINIAEF

RXN01694 - 5'-Region

CGACACCAATAATTTATTGGGTATCCACCAATTACCGCTGTGAGCACTGCAAATTACGTATTCGAAAAGCCATGTCCACCACG
TGTTCTATCCTGGCGGC

RXN01694 - kodierende Region

ATGCAAAAAATCACCCCAAACATCTGGTGCCAAGGCACCGCAGACGAAGCAGCCGAATTCTACGTCAATGCGTTTTCTGAGTT
TCCGGGTGGCGCAGAAGTACTCACCACAGTTAAGTATCCCGAAGCTGGCTTGCTGGACTTCCAGGAGCCTTTTCGCAGGAAAAA
CCTTGACGGTGGAACCTCGCTATCTCAGGCTTTAAGATCATCTTGATCAATGCTGGTGAAGAGTTCACTCCCAACCCATCGATC
AGCTTCATGGTGAATTTTGATGCGGTGCGTGATGAAAATGCCAAAGAGCACCTTGATGCGGTGTGGGAAAAACTCCATGAAGG
CGGCAGCACACTGATGCCAGTCGATACTTACCCATTTTCGGAATACTACGGGTGGGTACAAGACAAATATGGTGTGAGCTGGC
AATTGATGCTCAGCCGCCCAGAAGAAAAGCCAGGTCCCGCAGTAATCCCAACGCTCTTATTTGGTGGGGCAGCTCAAAATCAG
GCAGGCCCAGCTCAAGAAAACTACGTTGAGGTGTTCCCGAACTCCCAACTTGGTGATCGTGCACCTTATGGACAGCAAACAGG
TCTGCTGCACTCCTGAGGCCCTCATGTTTTCCAGTTCCAACTCGACGGTCAGTGGATTTTCGCGATGGATTCCGGAGTTGAGC
TAGATTTACCTTCAGTGAGGGTGCTCATTGATGTATGAAGCTCATGGTCAAGAAGAAGTTCGATGCCATCTGGAATGCACTC
TCGGCAGTTCCAGAAGCTGAGGCTTGTTGGTTGGTTGAAGGACAAGTTCGGCGTGAGCTGGCAGATTGTTCCCGACAACATGGA
GGAGCTCATGGCTAAACCCGGCGCGTATGAAAAGCTTCTTGCGATGAAGAAGATCAATATCGCGGAGTTC

RXN01694 - 3'-Region

TAGCAGTTCTAAGCGCTCCACGC

RXN01634 translatiert von RXN01634 (2339) von 1 bis 552

MAVGQETDLLNAMETLAGVPSDLVVTEILPDTTSYDNFEFFEVHNTGSAPVTIGEGETTFAYSFDDSDTSDRKALDLGGEVT
VDAGETIVVWIEYSSSTVDTAAFSEQDFRDFYGMDSARIFRATGQAGLANGGDRGIRVLYNGEVSGWSHYPSDSAQVKGID
FALPKVGEQASIASAHQN

RXN01634 - 5'-Region

ATTGGATCTAGGAAAAATTATGAACATCTCAAAGCTCGGACTGAGGATCGCGGTCACAGCAACGGCGACCACAACGGCGCTTG
CCCTCGGTGGCACGCAC

RXN01634 - kodierende Region

ATGGCAGTTGGTCAAGAACTGATCTGCTCAACGCTATGGAAACACTCGCTGGGGTTCCCTCAGACTTGGTGGTCACTGAAAT
CCTCCCCGATACCACCAGCTATGACAATTTTGAATCTTTGAAGTTCACAACACCGGCAGTGCACCCGTGACCATTGGGGAAG
GGGAGTACACCTTCGCCTATTCTTTGACGATTCCGCCGATACGTCCCGCGACAAGGCAGTGGATCTTGGCGGGGAAGTCACG
GTAGATGCAGGCGAAACCATTGTGGTGTGGATTGAGTACTCCAGTTCAACTGTTGATACCGCTGCGTTTAGTGAGCAAGACTT
CCGTGATTTCTACGGCATGGATTCTCAGCCCGCATCTTCCGAGCAACTGGCCAGGCGGGTCTCGCTAACGGTGGTGATCGTG
GCATCCGAGTTCTGTACAATGGTGAAGTTTCTGGTTGGTCCCACTACCCATCAGATAGCGCAGCGGTTCAAAGGGAATTGAC
TTCGCGCTGCCAAAAGTAGGGGAGCAGGCCAGCATTGCAAGTGCGCACCAAAAC

RXN01634 - 3'-Region

GATCCAACCTCCAGGATCAATTA

RXN01592 translatiert von RXN01592 (5848) von 1 bis 1302
MADNFDNRSDNDRSSDRTPRGDRGDRGGYRNSRGNDNRGNRQNRDGESRDRGGYLGDRRDNRSGEYRQRDDRRDDRRDNRSD
DRRGGYRSDRNFDDRNSNMRRDRRGDRSYSRNDRSDRGYRSDRYDRNDRDDNRDTRGGDRGDRRYDRRDRDDRRDDRR
GGQGQGRPGDRRHANRAGAGRDQQRDSLHPQRAGFREERLNTRLNEPDLPGDIDIKDLPLVLQDLKVLKDNADAVAKHMI
MAATWLADDPQLALRHARAADKDRAGRVSVVRETNGIAAYHAGEWKEALSELRAARRMSGGPGLIAMADRCERGLGRPEKAIEL
AREEDLSSLDQDNLIELAIVVAGARHDLGQHDSAIVELQKVNPSLKSTGTFHSRLSYAYADALVLAGRGDEAREWFQHAATLD
EDGYLDAERIEQLDNGNN

RXN01592 - 5'-Region

GATTCAATGGTAGAGGTATTTTCATTTCTCCCATTTGGATCACACTCTTTTTGTCGTTCTATAATAAAGTCTTTACCCGGTTCT
CAACATCAAGGAACAAC

RXN01592 - kodierende Region

ATGGCTGATAACTTCGACCGCTCTCGTGACAACGATCGCTCAAGCGACCGCACTCCTCGTGAGACCGCGGCGATCGCGGTGG
CTACCGAAACTCCCGCGGCAACGATGACCGCGGAACTACCGTCAAAACCGAGATGGTGAATCACGAGATCGCGGCGGATACC
TCGGTGACCGTCGCGACAACCGCTCAGGTGAATATCGTCAACGTGACGATAGGCGTGACGATCGTAGGGACAACCGAAGCGAT
GACCGCCGCGGTGGTTACCGCTCCGATCGCAACTTTGACGATCGCAACAGCAACATGCGTGATGATCGTCGCGGCGGCGACCG
TTCATACAGCCGTAATGATCGCTCCGATCGTGGCTATCGTAGCAATGACCGCTACGACCGTAATGATCGTCGCGATGACAACA
GGGACACCAGGGGTGGCGATCGCGGAGACCGTCGCTACGACAGGCGCGATGACCGACGTGATGATCGTCGCGATGATCGTCGT
CGCGGACAGGGCCAGGGGCGTCCAGGTGGAGATCGTCGACATGCGAACCCTGCGAGGTGCGAGGTGCGGATCAGCAGCGTGATTC
GCTGCATCCACAGCGCGCTGGTTTCCGTGAAGAGCGCTTGAACACTCGTCTCAATGAGCCTGATTTGCCGGGCGATATTGATA
TCAAGGACTTGGATCCTTTGGTCTGCAGGATTTGAAGGTTTTGTCTAAGGACAACGAGATGCCGTGCGAAAGCATATGATC
ATGGCTGCGACCTGGCTGGCCGACGATCCTCAATTGGCACTGCGCCACGCCCCTGCTGCGAAGGATCGCGCGGGACGCGTGTC
CGTCGTGCGTGAGACGAACGGCATTGCTGCTTACCATGCTGGTGAGTGGAAGGAAGCTCTTTCCGAGCTTCGTGCTGCGCGCC
GCATGTCGGGTGGTCTGGTTTGATTGCTGTGATGGCTGACTGTGAGCGTGGCTTGGGTGCTCCTGAGAAGGCGATTGAGCTG
GCTCGAGAAGAGGATTTGAGCTCTTTGGATCAGGATAACCTGATCGAGTTAGCGATTGTTGTTGCTGGAGCACGCCATGATTT
GGGTCAGCATGACTCTGCGATTGTGGAATTGCAGAAGGTTAATCCAAGCTTGAAGAGCACCGGTTTACCCATTCTCGTTTGT
CTTACGCTTACGCCGATGCGCTTGTGTTTGGCTGGTCGTGCGGATGAAGCACGTGAGTGGTTCCAGCACGCTGCCACCTTGGAT
GAGGACGGCTACCTCGATGCAGAGGAGCGTATCGAGCAGCTCGATAATGGGAACAAC

RXN01592 - 3'-Region

TAGACTATTGGTCTAGAGTGTA

RXN01574 translatiert von RXN01574 (7504) von 1 bis 1035

MSNKRIGVVIVSYGHEQDVANLVDTFADQLKTGDRVVVDNRKPWVLKDAVGERLEKHGAEIINHNDGGFAAGCNVGAAHIVD
DVDLLFFLNPDFTVIDDPTLFNSLRRVDEQWAAFMPYLLLPDSTINSAGNALHISGLSWVTGLDEKPVEGSSEVTDISIASGAC
LAVRVDWWKRLGGMEELYFMYHEDTDFSARLLLGGGRIGLLHSAYVTHHYDYAKGDYKWIYIERNRHVLLLSVLPLPLLFVLI
PQILGVNLGLWAIAAKEKRVGLKVKSRLRLIIRDLPALFKLRRSTQELAEIPTSQYLAKMEWRLDNPNLGNIGSNKIVATGYKT
YYKLCMSILKLLA

RXN01574 - 5'-Region

AAAAATCCGCACACCCACATTGCTCAAAGGAATTGACAAGGCGCTTCGCAAGGCCGAGTAAGGTGGGCAACGAGTCGGTGAG
AGTATGAGGAATATTTG

RXN01574 - kodierende Region

ATGAGCAACAAACGCATCGGTGTAGTGATCGTTTCCTACGGACACGAACAAGATGTTGCCAACCTGGTAGACACATTTGCAGA
TCAGCTGAAAACCTGGTGACCGCGTAGTTGTCGTGGACAACCGCAAACCTTGGGTGTTAAAAGACGCCGTGGGGGAGCGCCTGG
AAAAACACGGCGCAGAGATCATCAACCACGACAACGGTGGTTTCGCCGCTGGCTGCAACGTGGGCGCAGCGCACATCGTGGAT
GACGTTGACCTGCTGTTCTTCCTCAACCCCGACACAGTGATCGACGATCCCACCCTGTTCAATTCGCTGAGACGCGTCGATGA
ACAGTGGGCAGCATTTCATGCCGTACCTGCTGCTTCCTGACAGCACCATTAACTCCGCAGGCAACGCCCTGCATATTTCCGGAC
TGTCGTGGGTGACTGGTCTGGATGAAAAACCAAGTTGAAGGCTCATCTGAAGTTACCGATATTTCCATTGCCTCTGGCGCCTGC
CTTGCCGTGCGCGTGGACTGGTGGAACGCCCTCGGTGGCATGGAAGAACTGTATTTTCATGTACCACGAAGACACTGACTTCTC
GCCCCGCTTGCTGCTGGCCGGCGGTGCAATTGGTCTCCTGCATTCCGCGTATGTCACCCACCATTACGACTACGCCAAGGGTG
ACTACAAGTGGATTTACATCGAACGAAACCGACACGTTTTGCTGCTCAGCGTGCTGCCGCTTCATTGCTGTTTCGTGCTGATC
CCGCAGATCCTCGGTGTGAACCTGGGACTGTGGGCGATTGCCGCAAAGGAAAAGAGGGTCGGACTCAAGGTGAAGTCCCTTCG
CCTCCTGATCCGCGATCTACCAGCGATTTTCAAACCTGCGTAGGAGCACGCAGGAGCTTGCCGAACCTCACACCATCGCAATATC
TGGCAAAAATGGAATGGCGCCTAGACAATCCCAACCTAGGCAACATTGGATCCAACAAGATTGTTGCGACTGGATATAAGACC
TATTACAAGTTGTGTATGAGTATCCTGAAATTGCTCGCT

RXN01574 - 3'-Region

TAACACCCCATAAAGAGGGTGAA

RXN01492. translatiert von RXN01492 (9162) von 1 bis 804

VTTTLWAVSDLHAAVKANADPIENIQPKDPSDWLIVAGDVAERTELVLLEILARLRRRFKVIWVPGNHELFSRSADRYQGRDK
YSELVEGCRKIDVLTPEDPYLTFGGVTIVPLFTLYDYSFRRPGFTVEQAVQAARDRQVMMTDEFSIAPFVDIRAWCWDRLAYS
IKRLSKINGPTILINHWPLVVEPTYQMRWQELALWCGTRHTRGWAERYNAEAVIYGHLHMPGITNVNGVKHIEVSLGYPREWE
HWSGQHVWPYPVMEVDNAG

RXN01492 - 5'-Region

TGAAGTGGGCGCGTTCGCGCGCGTGGCAAACCACATTTGCTGGACTGCAGGCATAATGGACAAAGACTTTTAGCAACAA
CTATGTGGAGGTGACAG

RXN01492 - kodierende Region

GTGACCACAACCTCTGTGGGCGGTTTCTGACCTCCATGCAGCGGTGAAAGCTAACGCTGATCCAATTGAGAATATTCAGCCTAA
GGATCCGCTCTGACTGGTTGATTGTTGCAGGTGACGTGGCGGAACGTACCGAATTGGTGCTGGAAATTTTGGCACGTTTGCGCA
GGCGCTTTGCCAAGGTGATCTGGGTTCCGGGTAAACCATGAGTTGTTTTCTCGCTCCGCGGACCGCTATCAGGGGCGCGATAAG
TACTCTGAACCTCGTTGAAGGCTGCCGCAAGATTGATGTGTGACTCCGGAGGACCCCTACTTAACCTTTGGTGGGGTTACTAT
CGTTCCGTTGTTTACACTCTATGACTACTCGTTCCGTCGACCTGGTTTCACCGTGGAGCAGGCTGTGCAAGCGGCGCGGGATC
GTCAAGTGATGATGACTGATGAGTTTCTATTGCGCCTTTTGTGATATCCGAGCGTGGTGCTGGGATCGCTTAGCCTATTCC
ATTAAGCGTTTGAGCAAGATCAACGGGCCAACAAATTTGATTAACCACTGGCCGCTGGTGGTGGAACCGACTTATCAGATGCG
CTGGCAGGAACCTTGCATTGTGGTGCGGTACTCGCCACACCAGGGGATGGGCCGAACGCTACAACGCGGAAGCCGTTATTTACG
CATCTGCATATGCCCGGAATAACCAACGTCAACGGTGTGAAACACATTGAAGTGTCGTTGGGTTATCCGCGCAATGGGAG
CATTGGTCTGGGCAGCATGTGTGGCCATATCCAGTGATGGAGGTGGACAATGCTGGA

RXN01492 - 3'-Region

TGAGTCTTTGTTTCCAAATTCGG

RXN01485 translatiert von RXN01485 (9216) von 1 bis 2052

MNFSLVHLRENVRRVSVTVVAIGAGALLISGPFFTAHTAEATETVYVLAESPEFYQDNVTDYTGQISSSDITNIQAAIDDVKASE
QKVIFVFLSSFDGVDPETWTQQALQANGGNVLIYALAPEERQYGIQGGTQWTDALDAANNAAFQALSQEDWAGSALALAE
SVGSSSSSSSSGSSSSSDFSGAWLAAAGVGTVAAGGGIWAYSRSRKKKTSATLEDAREIDPRDNTNRLMQLPMETLEHLAQEEL
TSTDDSIIRRGKEELAIATSEFGPERTRSFNRMNHSTGTLQKAFEIQQRLNDSIPESEAERQSMLEIISSCGQADDALDAEA
QNFADMRNLLINAGSKLDALTQKSVDLRLRLPKAQETLAGLRTRYSAEVLESIDDNVDLASASLDEAEVLPQAYEIESMPAG
EQGGLIDAIRHIEHAITTADKLLAGVEHADENISTAKANVADLIQEISDEINEAGQLKQSAGADGARADWASLDDAVRAASAA
LITASADAEEKDPLGTYTELVDVDSALDTQLDTRLRATAADQARQLRVFDQQLQSARSQIQKAEDLISTRGRIVKSEARHLANA
QKLYAMAQQNRTRDTRAGIDYGRQA AVAAQRASKSAQNDITTYNNRHNSGGGTTGAIVTGMVINSILNSGRGGGFCGGGGFGG
GGGGFSGGGGGGGGGFRGGRF

RXN01485 - 5'-Region

CTCAGGGCTGGCTGCACTACTGCGGTTAAGAGTAAACTCTTGGCTAAAAATCTTCTCACGTTAACTAGTGTGCCAGCTGGACT
CGTCTAAGGTGGGGACC

RXN01485 - kodierende Region

ATGAACCTTTAGTTTGGTGCATCTTCGGGAAAACGTCCGCCGTGTATCGGTCACGTGTGGCAATTGGAGCTGGCGCGCTCCTTAT
CAGTGGACCATTTTTTACTGCTCACACTGCAGAACCCACAGAACATATGTCTTGGCTGAATCACCAGAATTTTACCAAGACA
ATGTCACTGACTACACCGGCCAGATTTCTCTCGTCCGATATCACCAACATTCAGGCTGCCATCGATGATGTAAGGCATCTGAA
AAAAGGTTATTTTCGTTGTTTCTTAAGCTCTTTTCGACGGAGTTGACCCTGAAACGTGGACGCAGCAAGCACTCCAAGCCAA
CGGCGCGGAAACGTCTTGATTTATGCACTCGCTCCCGAGGAACGGCAGTACGGCATCCAAGGTGGTACTCAATGGACCGACG
CTGAACCTCGACGCCGCCAACACGCTGCTTTCCAGGCACCTTTCCCAAGAAGATTGGGCAGGCTCTGCACTAGCGCTGGCAGAA
TCAGTTGGTTCTAGTTCTTCCAGCTCTTCCGGCTCGTCCAGCTCTTCCGATTTTTCCGGCGCTTGGTTGGCTGCTGCGGGCGT
TGGCACAGTGGCCGCTGGTGGCGGAATTTGGGCCTATTCGCCGAGCCGCAAGAAGAAAAACAAGTGCTGCAACCTTGAAGATG
CACGCGAAATCGACCCGCGCGATACCAACCGCCTCATGCAGCTTCCCATGGAACTCTCGAACACCTTGCCCAAGAAGAGCTC
ACCTCCACTGATGACTCCATCCGTCGCGGAAAAGAAGAGCTCGCTATCGCTACCTCCGAGTTCCGACCAGAGCGCACCCGCGAG
CTTCAACCGCGCCATGAACCACTCCACCGGCACCTTGCAAAAAGCCTTTGAGATTACGACAGCGCCTCAACGATTCTATCCAG
AAAACTTTGCGGATATGCGCAACCTGCTGATCAACGCGGGCAGCAAAATTGGATGCTCTACCCAAAAATCCGTGCACTTGGC
CACCCGCTTCCCAAGGCCCAAGAAACACTCGCTGGCCTGCGCACTCGCTACTCAGCAGAGGTCCTTGAAAGCATCGACGACA
ACGTCGACCTCGCCAGCGCTTCGCTCGACGAAGCAGAAGAAGTCTGCCACAGGCGTACGAGATAGAGTCCATGCCCCGAGGC
GAGCAGGGGGGGCTTATCGACGCGATCCGTACATCGAGCAGCCATCACTACCGCAGACAACTCCTCGCGGGCGCTCGAGCA
TGCCGATGAAAACATCTCCACAGCCAAAGCAAACGTTGCCGATCTGATCCAAGAAATCTCAGACGAAATCAACGAAGCCGGCC
AACTCAAACAAAGCGCAGGAGCCGACGGTGCCCGTGGCGACTGGGCCTCCCTCGACGATGCTGTCCGCGCAGCCAGCGCAGCA
CTAATCACCGCATCAGCAGACGCCGAAAAGGATCCGCTCGGAACCTACACAGAACTTGTGATGTGCACTCCGCCCTCGACAC
TCAACTTGACACACTTCGCGCCACCGCAGCTGATCAAGCCCGCCAGCTACGCGTATTTCGACCAACAGCTGCAATCTGCAAGAA
GCCAAATCCAAAAGGCCGAAGACCTCATCTCCACCCGCGTTCGATCGTAAATCCGAAGCCCGCACCCACCTGGCCAAACGCA
CAAAAGCTGTACGCCATGGCACAACAAAACCGCACCCCGTGCAGGAATTGATTACGGACGTCAAGCAGCAGTCGC
AGCCCAACGCGCCAGCAAGTCAGCACAAAACGACATCACACCTACAACAATCGCCACAATTCCGGTGGCGGAACCACCGGTG
EATTGTCACCGGCATGGTGATCAACTCGATTCTAACAGCGGCCGCGGTGGCGGTTTCTGTGGAGGCGGAGGCTTTGGTGGA
CGCGGTGGCGGCTTCAGCGGTGGTGGCGGTGGCGGAGGAGGTTCCGCGGAGGCCGCTTC

RXN01485 - 3'-Region

TAGCCTGAAACGAGCAAAACCAA

RXN01471 translatiert von RXN01471 (6310) von 1 bis 807
MQSFKTLESWPVDNVSASVISDGAHFYGDVDRVFELMSVTKLLATYGFLVAIEEGVFELDSPMGPEGSTVRHLLSHASGVAF
DKPVAEKGVGERRIYSSAGMDILADAVAAEAEMPFAEYLREAVFEPLGMENSELWGSAGHEARSTVADLTKEFGQELTAPTLLIS
PETLAFAFQVQFPELIGTVPGYGMQKPCPWGLGFEIKGQKSPHWTGDLMPENTAGHFGQSGTFFWTVPGSGQVGVVLTDRNFG
PWAKPLWTAFNDEVWAEELNS

RXN01471 - 5'-Region
CCTCCAACCAGAGCTATAAAAAAGGTTGTAGAAGAAATTTCTTAACATACATCGTTGTGTGCCAGGGGTTTAACAGATTAGT
TGCTAAGGCTGTCTCC

RXN01471 - kodierende Region
ATGCAGAGTTTTAAGACGCTTGAATCTTGGCCTGTCGACAATGTTTCGGCCAGTGTCATCTCCGACGGCGCCGTCACACTTTTA
CGGCGACGTCGATCGTGTGTTTTGAATCATGAGTGTGACAAAACCTCGGCCACCTACGGTTTCCTTGTTGCCATCGAAGAAG
GTGTTTTTGAGCTCGATTACCGATGGGTCCTGAAGGATCCACAGTGAGGCATCTGCTCTCACATGCTTCCGGCGTGCGCATTC
GATAAGCCGGTAGCGGAAAAGGGAGTGGGGGAGCGCCGATTTACTCCTCTGCTGGCATGGACATCTTGGCGGATGCTGTTGC
CGCTGAAGCTGAAATGCCGTTTGCAGAGTATCTCCGCGAGGCTGTGTTTCGAGCCTTTGGGAATGGAGAACTCTGAATATGGG
GTTCTGCGGGCCACGAGGCGCGCAGCACAGTGGCTGATCTGACCAAGTTCGGCCAAGAGCTCACCGCACCAACTCTGATCTCA
CCAGAAACCCTTGCAGAGGCTTTCCAGGTGCAATTCGCGGAAGTATTGGCACCGTCCCCGGTTATGGCATGCAGAAGCCGTG
TSCGTGGGGATTGGGCTTTGAAATTAAGGGACAGAAGTCGCCGCACTGGACAGGTGACTTGATGCCGAGAACTGCTGGAC
TTTGACAGTCGGGAACATTCTTTGGACTGTTCCAGGCTCAGGTCAAGTCGGGGTTGTTTGGACTGACCGAAATTTCCGGT
CCTTGGGCTAAACCGTTGTGGACTGCCTTCAATGACGAAGTCTGGGCCGAGTTAAATTCA

RXN01471 - 3'-Region
TAAACTTTTGCCGTAATTACGGT

RXN01459 translatiert von RXN01459 (4910) von 1 bis 810
MNSNGFGFSFGNDDDDDKNRNNDPFGLFGGNFGFGGQGGAGGPGGLGDILNQFGQMLSGMGDSMNSPEAAGPVNYDLAARIA
RQOIGRVAPVKDSEKEAVEESLRLAELWLDNATQLPTSGHRVEAWNPNWLENTLPVWKRLVSPVAEQMNKAQLENLP EEARE
MMGPMSSLMNSMSSMNFGVQLGNALGDLAKQTLTGSDFGLPISPVGVS AVLPGNIAEASKGLNVAPQEMLVYICAREAAANAS
SSTCRGLSSASFPLRNTPSA

RXN01459 - 5'-Region

GATTCGGCGATTACGAATATCCACCCTACAACCTGTGTTCGCTTACAGCGTTGTTTCGGATCGTAGCAATAGGAAATAGTTGC
CTTGTAGATTGGTAAAT

RXN01459 - kodierende Region

ATGAACTCTAATGGCTTTGGTTTCTCTTTTGGCAACAACGATGACGACGATGACAAAAACCGCAACAACGACCCATTTCGGCCT
TTTCGGTGGCAACTTCGGATTTCGGAGGACAAGGTGGCGCTGGTGGCCCCGGCGGATTAGGCGATATTTTAAACCAGTTCGGCC
AGATGCTGTCCGAATGGGAGATTTCGATGAACTCCCCCGAAGCAGCAGGCCCCGGTTAATTATGATTTAGCAGCACGCATCGCC
CGCCAGCAGATTGGTCGCGTAGCGCCGGTGAAGGATTCTGAGAAGGAAGCAGTCGAAGAGTCCCTTCGCCTTGCTGAACTCTG
GCTCGATAACGCCACCCAGCTCCCCACTTCCGGGCACCGCGTCGAAGCGTGGAATCCAGAAAACCTCCCTGAAGAAGCCCGCGAG
CTGTGTGGAAGCGTCTTGTCTCCCTGTTGCGGAGCAAATGAACAAAGCCCAATTAGAAAACCTCCCTGAAGAAGCCCGCGAG
ATGATGGGCCCCGATGTCTTCTTTGATGAACTCCATGTCTTCGATGAACTTCGGAGTTCAATTAGGAAATGCGCTGGGCGACCT
GGCAAAGCAGACCCTCACCGGCTCTGACTTCGGTTTGCTTATCTCCCCAGTCGGCGTATCTGCCGTCTGCCCCGCAACATCG
GAAGCATCCAAAGGCCTCAACGTGGCACCAGCAAGAAATGCTCGTTTACATCTGTGCCCGCGAAGCCGCCCGCAACGCCTCT
TCAAGCACGTGCCGTGGCTTGTCGAGCGCTCGTTTCCTCCGTTGAGGAATACGCCGTCCGCC

RXN01459 - 3'-Region

TAGAAATCGACACCTCACACATT

RXN01409 translatiert von RXN01409 (6984) von 1 bis 822

WVDFPDVTARLGYEYMLTPHVDFGPFRRHPKADDDLVAALKKRAKDAGVTIPALLPVQRI SWPEETQORVAAVRNIRKRI
QLAVDLEVDTLNTEFSGRPERSEDAFYRSMEZLLPILEKEGIKFNIDPHPDDEFVENGIEAWRVIRGLNSKQVGFVYV
APHSFHMGDQAEAILPAVGDRLGAVYLSDTFDHEKSHGLRYITNPPGNAVRVHQHLKIGDGDVNFEEIFSLLRSTGYLDR
EDALLVSNVFAEDEAADEVSRYLEKIRSLIEN

RXN01409 - kodierende Region

TGGGTGGATTTCCCCGACGTCACCGCACGTTTGGGGTATGAATACATGCAGCTGACCCCTCATGTAGATTTCCGGTCCTTT
CTTCCGCCACCCCAAGGCAGACGATGATCTCGTGCAGCCCTGAAAAAGCGTGCCAAGGATGCCGGAGTCACCATTCTTG
CACTGTTGCCAGTGCAGCGTATTTCTTGGCCGGAGGAAACCCAGCGTGTTCAGCAGTACGCAACATCAAGCGCATCATC
CAGTTGGCCGTTGATCTGGAAGTAGACACCCCTCAACACGGAGTTTCTTGGACGCCCAGAACGCTCCGAGGATTCCGAAGA
TGCCTTCTACCGCTCCATGGAAGAACTCCTGCCATCTGGAAGAGAGGGCATCAAGTTCAACATCGACCCACACCCCTG
ATGATTTCTGTTGGAAGACGGTATTGAAGCATGGCGAGTCATCCGCGGTCTGAAGTCCAAGCAGGTGGGCTTTGTTTACGTG
GCACCTCACTCATTCCACATGGGTGATCAGGCTGAGGCAATCCTGCCAGCAGTAGGCGATCGCCTTGGGGCTGTGTACCT
GTCAGATACCTTCGACCACCACAAATCCACGGGCTGCGCTACATCACTAACCTCCAGGCAACGCAGTGCAGCGTGCACC
GCACCTAAAAATCGGTGATGGCGATGTGAAGTCTGAAGAGATCTTCTCACTGCTGCGCTCTACCGGTTACCTTGACCGT
GAAGATGCACTGTTGGTCTCCAACGTGTTGCAGAAGATGAAGCAGCAGATGAAGTATCCCGCTACCAGCTGGAGAAAAAT
CCGCTCACTCATCGAAAACGCA

RXN01409 - 3'-Region

TAGAGTTATCTCGAAACTACCAA

RXN01400 translatiert von RXN01400 (8655) von 1 bis 1500

MAISVSMPKAISREELERAWLEVIERHGTLRTVFSTGMGGEVQQHRIDVGPWKWIDHAVAPGESINEALRAVLNRQCSPPSRP
SHSLCIIDAHPRPTVIIIGSDHSHVDMWSMLVIVRDLAALDMELPVEPPLAFESHTAELLAAPPAPERIHQRWREILEAGGGK
MPQFPLPLGDAISMPERVEVRDIFGVNGLAIYSARARAQQVSSLALTISVMADVTAALADLPLRAVFPVHSRFDQRWHDVSGW
FISNSVIEVPDSDPHTAAQAVREAVSLGSYPLAELLEPWGGMPETPGMFAISWLDLRLPVSIDDIGLQAQYVSASLRTDGV
LWFILDRSGAHLRCRYPDSLVAENVGRWIDAIVAQMRAEAGTVNLQAGGEQLTLRHGTRADISEIAPLLARNAADPCELVDL
EHALDLLTHESSHFLAVVRNARGKIIAAMQLTIVPEFSRGGALHLHIEGPFIIPEYRTTDLDDKKLRAWAVEHGRARGVKVEEV
VG

RXN01400 - 5'-Region

GAGCAGCTATCACCTTATGCATCGGCTCCGGGTGAGGCGGTGCCGGTGTCTTTGATCAGCAACGACATGTGGGTGCCGGCG
AGCGACCGGGTAGTTGG

RXN01400 - kodierende Region

ATGGCGATTAGTGTCTTATGCCTAAGGCGATCAGCCGTGAGGAATTGGAGCGCGCGTGGCTTGAAGTGATTGAACGGCACGG
CACCTTGCGCACAGTATTTAGCACGGGAATGGGTGGGGAAGTGCAGCAACACCGCATTGATGTGGGCCCCGGGAAAATGGATTG
ACCACGCGCGTTGCGCCTGGTGAGAGCATTAAATGAGGCGTTCGCGGCGGTGTTGAATCGGCAGTGTTCCGCGTACTCGAGGCCA
TCGCATAGTTTGTGCATTATTGATGCGCATCCGCGTCCCACGGTGATTATTGGTAGTGATCATTCGCATGTGGATATGTGGTC
ATGCTGGTGATTGTGCGCGATTGTTGGCTGCGCTCGATATGGAACCTCCCGTTGAGCGCGCGTGGCGTTTGAATCGCACA
CGCGGAACTCCTCGCGGCTCCACCTGCACCGGAGCGGATTACCAGCGGTGGCGTGAAATTTTGAAGCTGGTGGTGGAAAA
ATGCCTCAGTTCCCGCTTCCGCTTGGCGATGCCATCTCCATGCCTGAACGCGTTGAAGTCCGCGACATTTTCGGGGTCAATGG
CTTGCGGATTACTCGGCACGCGCCCGCGCACAACAGGTAAGCTCCCTGGCCTTGACTATTTCACTGATGGCTGATGTTACGG
CGCGCGTGGCGGACCTTCCGTTGCGTGCTGTATTCCCAGTCCATAGCCGTTTCGATCAGCGCTGGCAGGATAGTGTGGGTGG
TTTATCTCCAATTCGGTCATTGAGGTGCCAGATTCCGATCCTCACACTGCAGCCCAAGCTGTTCCGCGAGGCTGTTTCTTTGGG
AAGCTACCCGCTGGCTGAGCTGCTTGAACCGTGGGGTGGCATGCCGGAACGCCAGGAATGTTTGCTATTTCTTGGCTTGACC
TGCGCCGACTCCCAGTGAGCATTGACGATATTGGCCTCCAAGCCAGTATGTCAGCGCTTCACTGCGCACCGATGGTGTGATG
CTGTGGTTTATTTTGGATCGCTCCGCGCGCACCTTCGCTGTCGTTATCTGACTCTTTGGTGGCGCGGGGAAAATGTAGGCCG
CTGGATTGATGCGATTGTTGCTCAGATGCGCGCCGAAGCTGGGACGGTGAATCTGCAGGCCGCGGGGGAACAGCTGACACTTC
GGCATGGAACTCGCGCGCGATATTTCCGAGATCGCCCCGCTACTTGCCCCGAAACGCAGCTGACCTTGTGAGCTGGTGGATCTT
GAACATGCCCTTGACCTGCTGACACACGAGTCTTCGCATTTCTGGCGGTGGTTCGAAACGCTCGTGGCAAGATAATCGCAGC
GATGCAGTTGACTATTGTTCCGGAGTTTCCCGCGGTGGTGGCTTCATCTTCACATCGAGGGGCGGTTTATTATCCGGAAT
ATCGCACCACCGATTGATAAAAAAGCTGCGCGCTGGGCTGTGGAGCATGGGCGTGCGCGAGGGGTGAAGGTTGAGGAGGTG
GTGGGA

RXN01400 - 3'-Region

TAGTTGGTTTATTGGCGCCTCGT

RXN01379 translatiert von RXN01379 (3028) von 1 bis 942

MSAFDSILNVEEWISDHYLTND DAKGASF SKRV RERIK EWKTTEDATQQSGPLTRFSSNRLQLQHALSELDDATTAASLVASA
LGYGVPSARHAQRGSDTISYSSWVGNAGSVEFLAATPAESFEENFRSLPLEPVAVNDKPQDITA AKLVGQIFLSDTPPAFVVI
TAGKWVVLAE RETWPLGRHLAIDISLVVERNDTKAQGEMQOTVVALARENTERAADGTTWWEETIEQSREHAVKVS GELRSAV
RESIEILGNDVLTRYEAKELSTAEIDGGELAKQSLRYLYRILFLLFAEASPELEILPTGTPEYDE

RXN01379 - 5'-Region

CAGAAATTGAGTAATGCTCTCGAGCCAGACCGTGAAC TTATTAGGCCTTTGGCCGTCATTCTTCCGCAGCCCCGCAACTTTGAA
CACCGAGGTTTAACACA

RXN01379 - kodierende Region

ATGAGTGCATTTGATTCGATCCTAAATGTTGAAGAGTGGATCAGTGACCACTACCTCACCAATGACGATGCCAAAGGTGCCTC
ATTTTCCAAGCGGGTGCGCGAGCGCATTAAAGAATGGAAAACCAACCGAGGACGCAACCCAGCAGAGTGGCCCTTTAACTCGTT
TTTCCAGCAACCGCCTGCAGTTGCAGCATGCTCTTTCTGAGCTTGACGACGCCACCACCGCCGCCAGTTTAGTGGCCTCTGCA
CTGGGGTATGGTGTCCCCAGCGCGCGCCACGCGCAGCGCGGCTCCGACACAATATCCTATTCTCTTGGGTGGGAAATGCCGG
CAGTGTGGAATTTCTTGCAGCGACTCCCGCTGAAAGCTTTGAAGAGAACTCCGATCCCTTCCCCTTGAGCCAGTAGCGGTCA
ATGACAAGCCCCAGGATATCACCGCAGCCAAATTGGTGGGCCAGATTTTCCTTAGTGATACTCCCCCTGCTTTTGTGTTATC
ACCGCTGGTAAATGGGTGGTTTTAGCCGAGCGTGAAACCTGGCCTCTAGGCCGCCACCTAGCTATTGATATTTCCCTGGTGGT
GGAACGTAATGACACCAAAGCCCAGGGTGAGATGCAGCAGACGGTCGTAGCACTAGCCCGGAAAAATACCGAGCGTGCCGCCG
TGGCACCACTGGTGGGAAGAAACCATTGAGCAATCCCGCGAACATGCTGTCAAGGTTTCTGGCGAGCTACGCAGTGCCGGTG
CGTGAATCCATTGAAATCCTGGGCAATGACGTGCTCACACGCTATGAAGCTAAAGAGCTCTCCACCGCTGAGATCGACGGTGG
CGAGCTAGCTAAGCAATCTTTGCGCTATCTCTACCGCATTTTGTTCCTGCTTTTGGCCGAGGCTTCACCAGAGCTTGAAATCC
TGCCAACCGGCACCCCGGAATATGACGAG

RXN01362 translatiert von RXN01362 (8917) von 1 bis 3654
KWLGTPAQSLFLIPPVTOHARTGTHLYDSLQLLFTLVDKGHHPTDAKAVAFDAEAGEEGLHFRNLSADLFLPAATELIDR
VGLSNEALNKVLENLLLSRVQSGKDRGFI SYATLGVTELQVYEGLMSTYGFIAQEDLFEVAPHGKADKGSWMLPVSKAD
EVPADSFIEVDQEAPGGGVIKVRKRHPRGSFVFRQSSRDRERSASFYTPQVLTSTFTVTQAIEELQASKRITTANDVLSLT
ICEPAMGSGAFAVEAVRQLAELYLELRQEELEQQI PAEDRAKELQKVKAHIALHQVYGVLDNSTAVELAEI SLWLDTMNA
EMDAPWYGLHLRNGNSLVGATRSLYAPSLNKKAWLTATPTRYRLDDIAQAIDENKAEPLFNHGIHHFLLPSTGWGATAD
AKDLKDLMA TEIKELKSWRTSIRASLSKTQIKQLNNLALRVETLWRFVLMRIRIAESQISRSTTLWGQEPAEVSEVV TRE
QIEQDLFGNIDGAYNRLRLVMDAWCALWFWPLDAVATAEHPERPALPDLEWLATLTEILGIDLPLKSKNENQIVLGPDT
NWLAINDAEATDLGFSGALSFERVSANHPWINVARQVAKQQSFFHWDLDFAHVFAKGGFDLQVGNPPWVRPDVNFEDLLA
EHPDWAVMSKPTQASKKERQKNFNHNNPKSLEHVVS GAGEPVATSAILGSVTLYPHLKDQRPDLRGFMEXTWSNASPAG
AISLIHPESHFTEKKAAPLRRGAYERLRHRHWFINELILFDVHDLVKYGVHVGAPQESINFLSAASLYHPQTVLDSFDH
DGSGLNPLGLKDDNGNWD RRP HKDRIQLVNADTLTVWKSILEDEQTPYLDTRMVYTVNTEAAAALEKLASAPRIKELGLQF
SSGWNETTDKKKG YFDVWGYPASWSDAILQGPLGVATPMIKQPNPTMKHNQDWSEIDFEAI PANFI PATAYQPDROT K
PTYDADYGTWTFGDKQVPVADTFRIAWREMAATTGFRTPVPSVIPPGANHVHTVNSAASRSNLKTI LVGAQLGAILSDYF
ARSSGSSHIFNDIVRKIPLPNFTSLEKQFARTYLRLNCLTSAYAPLWEEITGEPWDVQVPLRNAEQRRAAQNDIDAMVAL
SLGISADELCMIYRTQFPVMRRYDQEDHFDANGRKVPKEIKLQQLKDKGQELSVEKRTWVHPQSEVSYTFEYPFRVLDR
EADLRAAYAKFENQLKEP

RXN01362 - kodierende Region

AAATGGCTTGGCAGCCTGCGCAGAGCTTATCTTAATCCCCCAGTTACCCAGCATGCTCGCACGGGCACCCACCTTTA
TGATTCCCTGCAGCTGCTGTTCACTCTGGTGGATAAAGGCCACCACCCAACAGATGCTAAGGCTGTAGCTTTTGATGCCG
AGGCTGGAGAAGAAGGCCTGCACTCCGCAACCTTTTACGCGGATCTCTTCTCCCTGCAGCCACAGAACTTATTGATCGA
GTTGGTCTTTTCCAATGAAGCCCCAAAACAGGCTTTGGAACCTCCTGCTCTCCCGGTGCAATCCGGTAAAGACCGCGG
CTTTATCTCTATGCCACCTTGGGTGTTACCGAGCTTGGCCAAGTTTATGAGGGTCTGATGTCTTATACCGGCTTTATCG
CCCAGGAAGATCTTTTGGAGTTGCACCACATGGCAAAGCCGATAAAGGTTCTTGGATGCTCCCGGTCTCAAAGGCTGAT
GAAGTCCCTGCCGATAGCTTTATCGAAGTTGATCAAGAAGCCCCCTGGTGGCGGCGTAATCAAGGTGCGTAAACGCCACCC
GCGCGGATCATTTGTGTTCCGTCAGTCTCTCGTGACCGCGAACGCTCAGCGTCTTCTACACCCACAAAGTACTCACCA
GCTTTACTGTCAACCCAGGCTATTGAAGAACTCCAGGCATCAAAGCGCATCACCAAGCCAAATGATGTTCTCAGCCTCACC
ATCTGTGAACCTGCCATGGGTTCCGGCGCTTCCGCTGTGGAAGCAGTACGCCAATTAGCAGAGCTTTATTTGGAATTGCG
CCAAGAAGAACTAGAGCAGCAGATTCCAGCGGAAGACCGTGCCAAGGAACCTCCAAAAGGTCAAAGCGCACATTGCGCTGC
ACCAGGTTTATGGTGTGGACCTTAACAGCACTGCTGTGGAGTTGGCGGAAATCTCGCTGTGGCTAGACACCATGAATGCA
GAAATGGACGCACCTTGGTATGGCCTGCACCTGCGTAATGGTAACTCCCTCGTTGGTGCCACCCGTTTCCGCTGTATGCACC
TAGTCTGCTTAATAAAAAAGCCTGGTTAACTGCTACTCCAACCCGCTATCGGCTTGATGATATCGCGCAGGCTATTGATG
AAAAAAGATCTTAAAGATCTTATGGCTACTGAAATCAAGGAGCTTAAATCTTGGCGTACTTCCATCCGTGCGTCTTTGAG
TAAACTCAGATTAAGCAGCTCAATAACCTTGCCCTACCGGTGGAACACTATGGCGATTTGTGCTGATGCGTATTTCGA
TTGCAGAATCCAGATCTCACGTAGCACTACTCTCTGGGGTCAAGAGCCAGCTGAGGTTTCGGAGGTTGTCACACGTGAG
CAAATTGAACAAGACCTCTTTGGCAATATTGATGGTGCATATAACCGTCTACGCTTGGTGATGGATGCTTGGTGTGCGCT
GTGGTTCTGGCCTTTGGATGCTGTTGCTACCGCTGAGCATCCGGAGCGTCCAGCCCTTCCAGATCTTGATGAGTGGCTAG
CCACCCTGACGGAGATTCTGGGTATTGATCTCCCTCTGAAGTCCAAAAACGAAATCAGATTGTCTTAGGTCCAGATACC
AATTGGCTAGCCATTAATGATGCCGAGGCTACTGATCTTGTTTTCTGGGGCATTGAGCTTTGAGCGTGTAGCGCGAA
TCACCCGTGGATCAATGTTGCCCCGCAAGTGCTAAACAACAGAGCTTCTTCCACTGGGATCTAGACTTCGECACGTTT
TTGCCAAGGGTGGATTTGATCTGCAGGTGGTAATCCACCATGGGTGCGACCAGATGTGAACCTTGAGGATCTGCTTGCT
GAACATGATCCGTGGTGGGCTGTAATGAGTAAACCAACCCAGGCATCCAAAAAGAACGCCAGAAGATTTTACACAACAA
TCCTAAGAGCCTCGAACATGTGGTCAGTGGTGCAGGTGAACCCGTGGCTACTTCTGCGATCTCGGTTCCGTTACTCTCT
ATCCGCATCTTAAAGATCAACGCCCGGACCTTACCGGGGCTTTATGGAAAAGACTTGGTCTAATGCCTCCCCGGCAGGT
GCGATCTCACTGATCCACCCCGAATCCCACTTCACTGAGAAAAAGCTGCCCCGCTACGTCGGGGAGCATATGAGCGTCT
GCGCCGCCACTGGCAGTTTCAATTAATGAGTTGATTCTTTTACGCTCCACGACTTGGTTAAATATGGCGTACATGTCTATG

GCGCTCCGCAGGAATCTATTAACCTTTTAAAGTGCTGCGTCGCTTTATCACCCACAAACAGTGCTTGATTTCATTTGATCAT
GACGGTTCAGGTAATCTCCCTGGTCTTAAAGACGACAATGGCAACTGGGACCGTCGCCCACACAAGGACCGTATCCAACCT
GGTCAATGCCGATACTTTGACGGTGTGGAAGTCCATCCTGGAGGATGAACAAACGCCATACTTGGATACCCGCATGGTTT
ATACCGTCAACACGGAAGCAGCAGCAGCGTTGGAAGAGTTGGCTTCTGCACCTCGTATCAAAGAACTCGGGCTGCAGTTC
TCCAGTGGCTGGAATGAAACCACCGATAAGAAAAAGGGATACTTTGACGTTGGTTGGGGCTACCCAGCTTCCTGGTCTGA
TGCCATTTTGCAGGGGCCGCACCTGGGTGTTGCTACACCAATGATCAAGCAGCCCAATCCGACAATGAAGCATAATCAAG
ATTGGTCTGAAATTGATTTGAGGCCATTCTCTGCAAACTTCATACCTGCAACGGCGTACCAGCCCGATCGCCAAACAAAG
CCCATTATGATGCTGACTACGGCACCTGGACTTTCGGGGACAAGCAGGTACCAGTTGCAGACACTTTCGGAATTGCATG
GAGGGAGATGGCTGCCACCACGGGATTTAGGACTGTCTACCCATCAGTAATTCACCGGGAGCCAACCATGTGCACACAG
TTAATAGCGCTGCATCAGTTCAAACCTAAAAACCATTCCTCGTTGGAGCACAGCTTGGTGCAATTCTAAGTGACTATTTT
GCTCGGTCTCGGGTTCAAGCCACATATTTAACGACATTGTTGCAAGATTCCACTTCCAAATTTACATCCTTGAAAA
GCAGTTCGCCCCGACATACCTCCGCCTCAACTGCCTGACCTCAGCTTATGCCCCATTGTGGGAAGAGATCACCGGTGAGC
CGTGGGATGTTTCAGGTGCCTTTGCGCAATGCCGAGCAACGTCGAGCAGCGCAAAACGATATTGATGCCATGGTGGCATTG
TCTTTGGGTATTAGTGCTGATGAGCTGTGCATGATTTATCGCACTCAATTCCCAGTGATGCGTAGATATGATCAAGAAGA
TCATTTTGATGCCAATGGCCGTAAAGTTCTTAAAGAGATCATCAAGCTGCAGCAGAAACTTAAAGATGGCCAAGAGCTCA
CCGTGGAAAAGCGCACCTGGGTGCATCCCCAATCAGAAGTGTCCTATACCTTTGAATATCCTTTCCGGGTGTTGGATCGT
GAAGCTGATCTGCGTGCTGCATATGCAAAATTTGAAAACAGCTTAAGGAGCCA

RXN01362 - 3'-Region

TAGAGCGCTTATGTCCTCACTCA

RXN01324 translatiert von RXN01324 (6882) von 1 bis 300
MSNSECHTHGYIEEKQRYLARKRIEGQTRGIHRMIDEEQYCIDILTQISAVNSALKNVAFGLDDHLAHCVKEAADLGGDEL
DAKLKEVSDAIARFSKA

RXN01324 - 5'-Region
TCCTCAACCCAATGCTTGCCGGCATTGCGATGGCCTTCAGTTCAGTTTTCGTCGTCTCCAATTCCTTGCGTCTGCGAGGATTC
AAAGCAAGGAGCAACTA

RXN01324 - kodierende Region
ATGTCCAACAGCGAATGCCACACCCACGGTTACATCGAAGAAAAGCAGCGTTACCTCGCACGCCTCAAAAGAATCGAAGGCCA
AACCCGAGGCATTACCCGCATGATCGACGAGGAACAATACTGCATCGACATCCTCACGCAGATCTCCGCAGTGAATCCGCAC
TCAAAAACGTGGCGTTCGGCCTCCTCGACGATCACCTCGCTCACTGTGTCAAAGAAGCAGCTGACCTCGGCGGCGACGAATC
GACGCAAAACTCAAAGAAGTTTCCGACGCCATCGCCCGCTTCAGTAAGGCC

RXN01324 - 3'-Region
TAAACGGATCCGGTGGCATTGGA

RXN01296 translatiert von RXN01296 (1595) von 1 bis 1083

MSIEQAITSLSARVRELKPIIETEEATKTALIIPFISNVLGYDVTDPREVIPEYTADVGVKKGEKVDFAIKTGDDFHFLIECK
KVGSPSLDHANQLVRYFNVTDTFAILTNGEIYQFYGQLDAANRMDAKPFMTLDLNNIDARQFPHEMCTRKHFNPAALAN
AEELKYIAELKKVIANQFQEPDVEIVKMLAATVTTKRMTAQNLEFFTRLVNTASSQFLKDEVNRRLSAQVFEDPVQTQGADA
ETPAEDEAVIEEVVSEIVTTEEEIHGHSIVRAICCSEVSAQEITMRDAKSYCAILFQDNNRKPIARFYFDRKIPRIGIFNAEG
EQEHFDLESIEDIYNHADLLHSRVVALNA

RXN01296 - 5'-Region

TAAAACTCCTGTTTCGAAATTGCCAGACAGGTGTCCACCAACTGCTACATTGCATTGGAGATCCAGAATCGATCACTCTAC
GACCAGGAAAACTTTC

RXN01296 - kodierende Region

ATGTCCATCGAGCAAGCAATCACTTCACTCTCCGCAAGAGTGCGGGAATAAGCCCCATCATCGAGACTGAAGAAGCCACCAA
AACCAGCACTGATCATCCCTTTATCAGCAACGTTCTCGGCTACGACGCTCACTGATCCTCGTGAAGTCATTCCGGAATACACTG
CTGATGTTGGCGTCAAAAAGGGTGAGAAGGTCGACTTCGCTATCAAAACCGGCGATGATTTCCACTTCCTCATCGAATGCAAA
AAGGTCGGCTCCCCACTCAGCCTCGATCAGCTAACCAGCTCGTCCGCTATTTCAATGTACAGACACCGAATTTGCCATTCT
CACCACCGGCGAAATCTACCAATTCTATGGACAACCTCGATGCAGCCAACCGCATGGATGCAAAACCATTTCATGACCTTGGATT
TGAACAATATTGATGCCCGTCAGTTCCCTCATTTGGAAATGTGTACCCGCAAGCATTTCAACCCACAAGCGCTAGCCGCCAAC
GCTGAAGAACTGAAGTACATTGCTGAATTGAAGAAAGTCATCGCGAATCAATTCCAAGAACCTGACGTAGAAATCGTCAAGAT
CTTGGCGGCGACAGTCACCACAAAGCGTATGACTGCACAAAATCTGGAATTCTTCACCCGCTTGGTTAATACTGCGTCTTCCC
AGTTCTCTCAAAGACGAGGTCAATCGTAGATTGCGCTCCGCCCAAGTCTTTGAGGATCCTGTCCAAACACAAGGTGCTGATGCA
GAAACACCAGCAGAGGACGAAGCAGTAATCGAAGAAGTGGTTTCAGAAATCGTGACGACAGAAGAGGAAATCCACGGTCATTC
AATTGTCCGTGCAATTTGCTGCTCAGAGGTATCGGCACAAGAAATCACCATGCGTGACGCAAAATCCTACTGCGCTATTCTCT
TCCAAGACAACAACCGAAAGCCAATCGCCCGTTTCTACTTTGATCGCAAGATTCCACGCATCGGCATCTTCAATGCTGAAGGC
GAGCAGGAACACTTTGATTTGGAATCCATCGAAGATATCTACAACCACGCTGATCTTCTGCATTCCCGCGTCGTAGCATTGAA
CGCT

RXN01296 - 3'-Region

TAAGTTCTGCCTTTAGTTCTGCA

RXN01281 translatiert von RXN01281 (193) von 1 bis 762

VGLSRSHYQLAAENTKSLVRMPMTVPQGNPAKKLATDLNRNQVVDELSAAVSRGQLTLEEFEDRSSKAWNARHLDLVELI
SDVNDNPYTLGQQFPGASYAPAAAYETTPAMPNVSDPVNIVRNRITGNPNNGSKMSVSFMGGTVRKGGWHVPNVHTSFAMMGG
NQIDLRFDAFLESDRIQINAYTFMGGIEIIVPEGVFVICDGMGIFGGFEQSVDKAGALNPAPLPSNAPTVHIKGLAFMGGVSVV
TKKNI

RXN01281 - 5'-Region

GGACAACCTCATCAAATGCAAACATGCGCCACCGCGTCCCGGAAAACCAGGGGCGATACCGCGCTTGAGTGCGAGGTGCCGTC
AATCCCAACTACAAGA

RXN01281 - kodierende Region

GTGGGACTCAGCCGTTACATTATCAGCTAGCCGAGAAAACACCAAATCCCTTGTTAGAAATGGTTCCCATGACTGTTCTCTCA
AGGAAACGAACCAGCAAAGAAGCTCGCCACTGACCTTAACCGAAACCAAGTCGTGGATGAACTTTCTGCAGCTGTTTCCCGAG
GTCAACTCACTTTGGAGGAATTTGAAGATCGCTCCTCCAAAGCATGGAATGCCCGCCACTTAGACACTCTTGATAGAGCTGATC
TCTGATGTGAACGACAATCCCTACACTCTGCTTGGTCAGCAATTTCCCGGCGCCTCCTATGCGCCGGCGGCCTACGAGACCAC
TCCCCCAGCGATGCCCAATGTGTGATCCTGTAAATATTGTCCGCAACAGGATTACCGGCAACCCGAATGGCTCCAAAATGT
CGGTCTCGTTTATGGGTGGCACCGTGCGCAAAGGTGGATGGCATGTGCCTAATGTTACACATCCTTCGCCATGATGGGCGGC
AATCAGATCGATTTGCGCGACGCCTTCCTGGAAAGCGACCGCATTGATCAACGCCTACACATTCATGGGTGGCATCGAGAT
TATTGTTCCCGAGGGTGTTTTTGTGATTTGTGATGGCATGGGCATTTTCGGCGGCTTCGAACAGTCTGTGGACAAGGCCGGTG
ACTCAATCCCGCGCGCCTGCCAAGCAACGCGCCACGGTCCACATCAAAGGCCTGGCGTTCATGGGCGGAGTCAGCGTAGTC
ACCAAGAAAAACATT

RXN01281 - 3'-Region

TAAAAAGCTTGTCGACGCGCCCC

MKRRTITIAALALTSTLVLSACADNTEGENDTTTTIATTSAPDTTETTGTATDPETETGAAGEVSAEHNDADIMFAQMMI PHHQ
 QAVEMSEILLAKDDIPA EVIEFTQGVIDAQGPEIDRMNTMLETWEEDPVTGDMGEMDHGGMSGMMSEEDMTALEDAQGTRGCP
 ALP

GGGAAGGAGGAGGAAGGCTGCTGTGAGCTCTGCAGCCCCCGTCCGCTGACTTTCAATGCTTCATGGACTCCCCGATACACAACC
TCCGAAAGGGAACCCCC

ATGAAGCGCACTATCACCATCGCCGCTCTCGCTTGACCTCCACCCTGGTTTTGTCCGCTGCGCAGATAA CACTGAGGGGAGA
AAACACCGACACCACGACCATCGCCACTACGTCCGCCCCGACACCACCGAAACGACCGGGGCCACCACGGATCCTGAGACAG
AGACGGGGGCGGCCGGAGAGGTCTCCGCCGAGCACAATGATGCGGACATCATGTTCCGCGCAGATGATGATCCCGCATCACCAA
CAGGCCGTGGAGATGAGTGAAATCCTCCTGGCCAAAGGACGATATCCCGGCCGAGGTTCATCGAGTTCACCCAGGGTGTTATCGA
TGCCCAGGGCCCCGGAGATCGACCGGATGAATACCATGCTCGAGACCTGGGAAGAAATCCGGTCACCGGTGATATGGGTGAGA
TGGACCATGGCGGGGATGAGTGGAAATGATGAGCGAGGAGGACATGACAGCCCTCGAGGACGCCCAGGGGCACCCGAGGCTGCCCG
GCTCTACCT

TGAGCAGATGACCGCCCACCATG

RXN01167 translatiert von RXN01167 (6639) von 1 bis 396

MADRVLRGSRMGAVSYETDRDHD LAPRLVKYKTADGEIYEV PFADDAEIPEEWMCKNGKLGILMEGEGVESKPVKPPRTHWD
MLRERRSIEELDVLLERIEALRKRRRNAAKLLKAQQEAEAEKAAEEV

RXN01167 - 5'-Region

GTGTCGTTTTATTTCACACTAGCAGGTAAACTGTGTGAGACGAATCTTTTGCGGCAACTTGGAGGCCGCATTCTAGTAGGC
AAAGGACTGATACCTAG

RXN01167 - kodierende Region

ATGGCAGATCGCGTTCTTCGTGGCAGCCGCATGGGCGCCGTGAGCTATGAGACGGATAGGGACCACGATCTGGCTCCTCGCCA
GCTCGTGAAGTACAAGACCGCAGACGGGAAATCTATGAGGTTCCCTTCGCTGATGATGCGGAAATCCCCGAGGAGTGGATGT
GCAAGAACGGTAAGCTAGGCATCCTCATGGAAGGTGAGGGAGTCGAGTCCAAGCCGGTCAAGCCTCCACGTACTCACTGGGAT
ATGTTGCGTGAGCGTCGCTCAATTGAAGAGCTGGATGTGCTGCTGGAAGAGCGCATCGAGGCACTTCGTAAGCGTCGTCGCAA
TGCAGCGAAACTGCTGAAGGCTCAGCAAGAGGCTGAAGAAGCAGAAAAGGCAGCTGAAGAGGTT

RXN01167 - 3'-Region

TAATCTTCCTGCCTAAGTTAGAA

RXN01153 translatiert von RXN01153 (8673) von 1 bis 543

PFLSPIVTSTHAVVAYSTARGFGEHRVRWDYAQESPLRDTRGFDLRRYHQAPVVDPHAIGVANVFVPNGARFYVDHAHPE
YSSPEVTNAWDAMVYDAAGDHILMQAVSDVASFTSQNRSVLDGHDPCPALKIYKNNVDGKGASYGFHENYLYSRETDFDV
LAQALIPFFVCRQVIIGAGRV

RXN01153 : kodierende Region

CCGTTTTTAAGCCCGATCGTTACTTCCACGCACGCGGTGGTGGCGTATTCCACCGCGCGTGGATTTGGTGAGCACCGGGT
GCGTTGGGACTATGCGCAAGAGTCCCCACTGCGCGATACTCGTGGCTTTGATCTGCGCCGATACCACCAGGCCCTGTGG
TGGATCCGCACGCCATTGGTGTGGCCAACGTGTTTGTGCCCAATGGTGCCAGGTTTTATGTCGATCACGCGCACCCGGAA
TACTCCTCCCCAGAGGTCACCAATGCGTGGGATGCCATGGTTTACGACGCCGCTGGTGACCACATCCTTATGCAGGCCGT
CTCTGATGTTGCGAGTTTCACCAGCCAGAATAGGTCTGTGTTGGACGGCCATGATCCGTGTCCAGCTTTGAAAATCTACA
AAAACAATGTCGACGGTAAGGGTGCTAGCTACGGGTTCCACGAGAATTACCTCTACTCACGTGAGACGGATTTTGATGTG
CTGGCTCAGGCATTGATCCCATTTTTTGTGTGCCGGCAGGTCATCATCGGTGCCGGACGTGTG

RXN01140 translatiert von RXN01140 (3026) von 1 bis 933

MALDTRGEEMRFRPRALSGAPDTGKDPGLLLLDGQQLTTLYHCFSGDGYVNTVDFRSKKVTRKFYIDVAKAVESPVMSDEAI
FSVDETGKIIISHFGPVIDGGITDLETALAHGCLPVSVLLDDNGTDFLFDLADMAGEGAREHAKRFQSQIVKTLVSYDIPMIRL
DRETAKGGIGSIFAQANSSGLQMDVFDLLTAVFAADESVETEFSLRDDWVRVERNLRQHSALDGIGSTEFLLTAVALLVSARKG
HASGYREDILNLTAEYIPAADDEMIKGFDEAAEFLRQRCIFESRPGSIHRADCSTGRDHPHA

RXN01140 - 5'-Region

TATCCAACCTTCCCGATTTCACCGGGATTACGCTTGGGATGTTGATCGCATCCGAAGCCTTATAACCACTGTTCTTCGTGGTT
TTCCAGTAGGTGTGCTG

RXN01140 - kodierende Region

ATGGCACTAGACACCCGCGGCGAGGAAATGCGTTTCCGGCCACGCGCGCTGTCCGGCGCCCCAGATACGGGCAAGGATCCAGG
TCTTTTGCTTCTCGACGGACAACAGCGCCTCACCACCCTTATCATTGCTTCAGTGGCGATGGCTATGTAAATACGGTGGACT
TCCGATCAAAGAAAGTGACCCGGAAGTTTTATATTGATGTTGCTAAGGCTGTTGAATCTCCGGTCATGTCCGATGAGGCTATT
TTTTAGTTCGACGAAACCGGCAAAATCATCTCCCACTTCGTCAGTGATCGACGGCGGCATCACCATTAGAAAACAGCACT
TGCTCATGGTTGCCTTCCAGTTTCTGTGCTGCTGGATGATAACGGCACTGATTTCTCTTTGACCTCGCCGATATGGCAGGAG
AAGGCGCTCGCGAACACGCGAAGCGCTTCCAATCACAAATCGTTAAGACCTTAGTTAGTTACGACATCCCAATGATCCGACTG
GATCGTGAAACCGCCAAGGGTGGAATTGTTCCATCTTGCTCAGGCCAATAGCTCTGGCTTGCAGATGGATGTCTTTGATTT
GCTCACC GCGGTGTTTCGACGCCGATGAATCGGTGGAGACCGAATTCTCACTGCGTGATGACTGGGTGCGGGTTGAACGAAACC
TCGCCAACACTCCGCACTTGATGGCATCGGCAGCACGGAGTTCCTCACCAGTAGCCCTGTTGGTCAGTGCCCGCAAGGGA
CATGCGTCTGGTTACCGTGAAGATATCTTGAACCTGACGCTGGCTGAATACATTCTGCTGCCGATGAAATGATCAAGGGCTT
CGACGAGGCGGCTGAATTCCTGCGTCAGCGATGCATTTTTGAGTCTCGACCAGGTTCCATACACCGCGCAGATTGTTCCACTG
GCCGTGATCCTCACCCTGCT

RXN01140 - 3'-Region

TGATGCAGAAGAAATGGCCACCG

RXN01134 translatiert von RXN01134 (3892) von 1 bis 771

LSLDQVPYTAQIVPLAVILTLDDAEEMATARSWDRLNQWFWSGVLGELYGSPAVIARSGRDTDQVAAWIREGAGETAVVPKTI
RDTVFHESRLLSATQDTGVWKGIFALLMGRGARDWRTGQQFDRWTFDELGCNFHQIFPTKWCKERGIDPVLTESVLNRTPMGR
RTEVIGDTPPSRYLSRVQSKSLMGDEEFDQMLDTHLLSAEDLHSSNTTHFFASRRTNFIDMVEDAIGKAVIRDVNESDLTGG
HDGPSVHG

RXN01134 - 5'-Region

GATATCTTGAACCTTGACGCTGGCTGAATACATTCTGCTGCCGATGAAATGATCAAGGGCTTCGACGAGGCGGCTGAATTCCT
GCGTCAGCGATGCATTT

RXN01134 - kodierende Region

TTGAGTCTCGACCAGGTTCCATACACCGCGCAGATTGTTCCACTGGCCGTGATCCTCACCTGCTTGATGCAGAAGAAATGGC
CACCGCACGTTTCATGGGATCGTTTGAATCAGTGGTTCTGGAGCGGTGTCTTGGGTGAGCTTTATGGCTCGCCTGCTGTGATCG
CACGTTCCGGTCGCGATACCGATCAAGTTGCTGCGTGGATTCTGTGAAGGTGCCGGCGAAACTGCGGTTGTGCCAAAGACTATT
CGCGATACCGTTTTCCACGAATCACGTCTTCTCAGTGCAACCAAGACACCGGAGTGTGGAAGGGTATCTTCGCACTGCTCAT
GGGCAGGGGAGCACGCGACTGGCGCACCGGACAGCAGTTTGATCGCTGGACTTTTGATGAGCTGGGTGCAATTTCCACCAGA
TTTTCCCCACCAATGGTGCAAGGAACGCGGCATCGATCCTGTCCTTACGGAATCAGTACTGAACCGCACCCCGATGGGTGCG
CGTACCGAAGTAGTCATCGGCGATACCCCTCCATCCCGCTACCTGTCTCGTGTTCAATCCAAGTCGCTCATGGGTGATGAGGA
ATTTCGATCAGATGCTGGATACCCACCTGCTCAGCGCGGAAGACCTGCACAGTTCCAACACCACGCACTTCTTTGCTTCTCGAC
CCACCAACTTCATCGACATGGTCGAGGATGCCATCGGTAAAGCGGTGATCAGGGACGTCAACGAGTCAGATCTCACCGGTGGA
CACGATGGTCCCTCGGTACACGGG

RXN01075 translatiert von RXN01075 (2212) von 1 bis 411
MDNPFVNILNEQEALERLQSVSLGRVVRRSDEMDIFPVNFIVDKGAIYIRTAEGNKLFSMNLNHDVLFEADEVKDGKAWSVVV
RATAEIVRKLDEIATADTLELKPWIPTLKS NFARIVPNEITGREFTLGEEPERY

RXN01075 - 5'-Region

GCGGTGCGTGTTCATTAGCAGGTCAAGCGCGTAATTGAGGGCTAGACTGGTTAGTACCGGATATTCTTTTTTCTTTAGTTTG
TGGGAGTGGAGATAACT

RXN01075 - kodierende Region

ATGGACAATCCAGTCAACATCCTCAATGAGCAGGAAGCTTTGGAGCGCCTGCAGTCGGTGTCTCTTGGTCGCGTGGTGGTTTCG
TCGCGAGCGATGAGATGGACATTTTCCCGGTGAACTTCATTGTGGATAAGGGCGCAATTTACATTCGTACAGCTGAGGGCAACA
AGTTGTTTCAGCATGAATCTCAACCACGATGTGCTCTTTGAAGCCGATGAGGTCAAGGACGGAAGGCTGGTCCGTGGTGGTT
CGTGCGACCGCAGAGATTGTGCGCAAGCTGGATGAGATCGCTACTGCCGACACTTTGGAGTTGAAGCCTTGGATTCCAACCTT
GAAGTCCAACCTTTGCCCGTATTGTTCCGAATGAAATCACTGGGCGGGAGTTCACCCTCGGCGAGGAGCCTGAGCGCTAC

RXN01075 - 3'-Region

TAGCTTTGCCACATTTCACTAAA

RXN00946 translatiert von RXN00946 (9850) von 1 bis 774

MTHTLQATNPLDQTAWHAWHFSRNKEAISRTGATSLSATEWISATTLKDAHTFPSLPGRWYKRGSGVGAHLPPAFATTGTVQ
LRPGELLIAEDFTLTVIERLGQFALQVFDARNPKRFEFHSIAAFPPSEEWRIEARFFPEPDTVNTAAADGVIVATPTAGWVHF
LKGRLDYRLRVTVQKNNLRALFSDNSSTLG VYQHRFVDIPRPDAEGNTIIDFNRAYLPPKALNRKFLCPSPSLNNHLNLTVEA
GEKWVVAGG

RXN00946 - 5'-Region

GGGCCGGTGTGAATAAAACACCTTCCCCAAATAGACAGCATGGTCTAGATTAGCTTGAAACGAAAGCGTACATTTCGCAGCAAC
TAACGGAAAGCACACTC

RXN00946 - kodierende Region

ATGACTCACACTCTGCAGGCAACTAATCCCCTTGATCAAACCGCTTGGCACGCTTGGCATTCTCTCCGAAACAAAGAGGCCAT
CAGCCGCACCGGCCGCCACCAGCCTGAGTGCCACAGAGTGGATTAGCGCCACCACACTCAAGGACGCGCACACTTTTCCTTCAC
TTCCCGGGCGATGGTATAAACGAGGCGGCGGGGTAGTGGGAGCACATTTACCACCAGCTTTCGCAACAACCTGGAACGGTGCAA
CTGCGCCCCGGTGAGCTCTTGATAGCGGAAGATTTACCCCTCACCGTCATTGAACGGCTTGGACAGTTTGCACTTCAGGTTTT
TGATGCACGCAATCCGAAGCGTTTTGAATCCACTCCATCGCAGCTTTCCACCGTCCGAGGAATGGCGGATTGAGGCTCGCT
TCTTCCCGGAACCTGACACTGTTAACACCGCTGCAGCTGATGGAGTTATCGTTGCCACGCCCTACTGCGGGTTGGGTGCATTT
TTGAAGGGCCGTCTGGATTACCGTCTTTCGTGTAACCGTTCAGAAAAATAATCTACGGGCACTATTTAGCGACAATTCCTCGAC
GTTGGGCGTTTATCAGCATCGTTTTGTCGACATCCCACGCCCTGATGCCGAGGGAAACACCATCATTGATTTCAACCGCGCTT
TCTTCCCCCAAAGGCATTGAACCGAAAGTTCCTGTGCCCATCGCCAGCCTGAACAACCACCTCAATCTCACCGTGGAGGCA
GGGAGAAGTGGGTGGTTGCTGGAGGA

RXN00946 - 3'-Region

TAATACTTGCTAACCGTCTCTAAA

RXN00869 translatiert von RXN00869 (9381) von 1 bis 921

MTPPIISPESFEALRRMRAEPTMVAERFKQRRKRELLGEDGKLFIVAADHPARGALAVGNETAMANRYELLERMAIALSRP
GVDGVLGTPDIIDDLAALGLLDDKIVVGSMNRGGLRGASFEMDDRYTGYNVSSMVDRGVDFAKTLVRINLSdagTAPTLEATA
HAVNEAAAAQLPIMLEPFMSNWVNGKVNDLSTDAVIQSVAIAAGLGNDSSYTWMKLPVVEEMERVMEStTMPTLLLGGEGGN
DPDATFASWEHALTLPGVRGLTVGRTLTPQDGDVAAVDTAARLVHTDIIQFTSQSI

RXN00869 - 5'-Region

GCCAACACCGCGGGTGCGCTTGTGGCGTCCCGTCTTGAATGCTCCACCGCAATGCCTACTACCGATGAGGTGGAAGCCTCCCT
CAACCAGAAAGTCTGAT

RXN00869 - kodierende Region

ATGACTCCTCCGATTATCTCCCCAGAGAGCTTTGAAGCCCTAAGGCGGATGCGTGCGGCTGAACCCACGATGGTGGCGGAACG
TTTCAAGCAGCGCCGTAAGCGTGAAGTCTCGGTGAGGACGGCAAGCTGTTTATCGTGGCTGCCGACCACCCAGCGCGCGCGC
CACTTGCTGTTGGCGACAATGAAACCGCCATGGCTAACCGCTATGAACTGCTCGAACGCATGGCTATCGCACTGTCTCGCCCG
GGTGTGGATGGTGTGCTGGGAACTCCAGACATCATTTGATGATCTGGCGGCGCTCGGACTGCTCGATGACAAGATCGTGGTTGG
CTCCATGAACCGTGGTGGCCTGCGTGGCGCTTCCTTTGAAATGGATGATCGCTACACCGGCTACAACGTGTCTCTCCATGGTTG
ATCGTGGCGTGGATTTTCGCGAAAACCTAGTGCGCATCAACTTGAGCGACGCCGGAACCGCCCCGACCTTGAAGCCACCGCG
CATGCAGTCAATGAGGCTGCAGCAGCACAGCTGCCCATCATGCTCGAGCCGTTTCATGAGTAACGGGTAAACGGCAAGGTGGT
CAATGATCTTTCCACCGATGCAGTTATCCAATCTGTGCGCCATTGCTGCTGGTCTGGGCAATGATTCTTCCTATACCTGGATGA
GCTTCCAGTGGTGGAGGAGATGGAGCGCGTCATGGAATCCACCACCATGCCAACCTGTTGTTGGGCGGCGAAGGCGGCAAC
GATCCAGATGCCACCTTCGCATCCTGGGAGCATGCACTCACCTGCCGGGTGTGCGTGGCCTGACCGTGGGACGCACTCTGCT
GTATCCGCAAGACGGCGATGTGCGCCCGCTGTTGATACCGCAGCGCGACTTGTTCACACAGATATTCAACAATTCACTTCGC
AGAGCATT

RXN00869 - 3'-Region

TAAGGAATTTACACACATGTCTG

RXN00850 translatiert von RXN00850 (9820) von 1 bis 828

MANPFSKAWKYLMAFDSKIEENADPKVQIQQAIEDAQRQHQELSQQAAVIGNQRQLEMQLNRRRLAEIEKLQGNTRQAIQLA
DKARADGDVKKATEYENAAEFAAQLVTAEQSVEDTKQLHDQALQQADQAKKAVERNMSMALQQKVAERTKLLSQLEQAKMQEK
VSESLKSMDSLTSGSTPNLDQVREKIERRYANALGQAELASNSVEGRMAEVEQAGVQMGHSRLEQIRAEMAGGSLTAGNKQE
SIEAPAAGNNVTDDAVAQRMRELGEA

RXN00850 - 5'-Region

GGCAGTTGTGCCGGTATACGTGGCGTCATGGATAATGATCGCATAAATGAATACAGCGAAATTACCTGTTTCAGATATTTAGG
AAGGCTCACTTTTCAAC

RXN00850 - kodierende Region

ATGGCTAATCCATTCTCCAAGGCATGGAAGTACCTCATGGCGTTGTTTCGACTCCAAGATTGAGGAGAACGCGGATCCTAAGGT
ACAGATCCAGCAAGCCATCGAAGATGCTCAGCGCCAGCATCAGGAGCTCTCCCAGCAGGCAGCAGCTGTTATTGGTAACCAGC
GTCAGCTTGAAATGCAGCTGAACCGCCGTCTGGCTGAAATTGAGAAGCTGCAGGGCAACACCCGCCAGGCTATCCAGCTGGCT
GACAAGGCTCGCGCTGACGGTGATGTCAAGAAGGCTACTGAGTACGAAAACGCCGCTGAGGCTTTCGCTGCACAGCTGGTTAC
TGCTGAGCAGTCCGTTGAAGATACCAAGCAGCTCCACGACCAGGCTCTGCAGCAGGCTGATCAGGCTAAGAAGGCTGTGGAGC
GTAACCTCCATGGCTTTGCAGCAGAAGGTTGCTGAGCGCACCAAGCTTCTGAGCCAGCTGGAGCAGGCGAAGATGCAGGAAAAG
GTTTCCGAGTCCCTGAAGTCCATGGATTCTTTGACCTCCGGCAGCACTCCTAACCTGGATCAAGTTCGTGAGAAGATTGAGCG
TCGTTACGCTAACGCGCTTGGCCAGGCTGAGCTTGCGTCCAACCTCTGTTGAGGGCCGCATGGCTGAGGTTGAGCAGGCTGGCG
TTCAGATGGCTGGACACTCCCGCCTTGAGCAGATCCGCGCTGAGATGGCTGGTGGTTCCCTGACCGCTGGTAACAAGCAGGAG
TCCATTGAGGCTCCTGCAGCGGGCAACAACGTCACTGATGACGCAGTTGCACAGCGCATGCGTGAGCTGCGCGGCGAGGCT

RXN00850 - 3'-Region

TAAACCTGCCACAGCAGACTAAA

Patentansprüche

1. Isoliertes Nukleinsäuremolekül aus *Corynebacterium glutamicum*, das ein MCP-Protein oder einen Abschnitt davon codiert.
5
2. Isoliertes Nukleinsäuremolekül nach Anspruch 1, wobei das Nukleinsäuremolekül ein MCP-Protein codiert, das an der Produktion einer Feinchemikalie beteiligt ist.
- 10 3. Isoliertes Nukleinsäuremolekül aus *Corynebacterium glutamicum*, ausgewählt aus der Gruppe, bestehend aus den in Anhang A aufgeführten Sequenzen, oder einem Abschnitt davon.
- 15 4. Isoliertes Nukleinsäuremolekül, das eine Polypeptidsequenz codiert, ausgewählt aus der Gruppe, bestehend aus den in Anhang B angegebenen Sequenzen.
- 20 5. Isoliertes Nukleinsäuremolekül, das eine natürlich vorkommende allelische Variante eines Polypeptides codiert, ausgewählt aus der Gruppe von Aminosäuresequenzen, bestehend aus den in Anhang B aufgeführten Sequenzen.
- 25 6. Isoliertes Nukleinsäuremolekül, umfassend eine Nukleotidsequenz, die zu mindestens 50% zu einer Nukleotidsequenz, ausgewählt aus der Gruppe, bestehend aus den in Anhang A angegebenen Sequenzen, oder einem Abschnitt davon homolog ist.
- 30 7. Isoliertes Nukleinsäuremolekül, umfassend ein Fragment mit mindestens 15 Nukleotiden einer Nukleinsäure, umfassend eine Nukleotidsequenz, ausgewählt aus der Gruppe, bestehend aus den in Anhang A angegebenen Sequenzen.
- 35 8. Isoliertes Nukleinsäuremolekül, das unter stringenten Bedingungen an ein Nukleinsäuremolekül nach einem der Ansprüche 1-7 hybridisiert.
- 40 9. Isoliertes Nukleinsäuremolekül, umfassend das Nukleinsäuremolekül nach einem der Ansprüche 1-8 oder einen Abschnitt davon und eine Nukleotidsequenz, die ein heterologes Polypeptid codiert.

10. Vektor, umfassend das Nukleinsäuremolekül nach einem der Ansprüche 1-9.
11. Vektor nach Anspruch 10, welcher ein Expressionsvektor ist.
- 5 12. Wirtszelle, die mit dem Expressionsvektor nach Anspruch 11 transfiziert ist.
- 10 13. Wirtszelle nach Anspruch 12, wobei die Zelle ein Mikroorganismus ist.
14. Wirtszelle nach Anspruch 13, wobei die Zelle zur Gattung *Corynebacterium* oder *Brevibacterium* gehört.
- 15 15. Wirtszelle nach Anspruch 12, wobei die Expression des Nukleinsäuremoleküls eine Modulation der Produktion einer Feinchemikalie von der Zelle bewirkt.
- 20 16. Wirtszelle nach Anspruch 15, wobei die Feinchemikalie ausgewählt ist aus der Gruppe, bestehend aus organischen Säuren, proteinogenen und nichtproteinogenen Aminosäuren, Purin- und Pyrimidinbasen, Nukleosiden, Nukleotiden, Lipiden, gesättigten und ungesättigten Fettsäuren, Diolen, Kohlehydraten, aromatischen Verbindungen, Vitaminen, Cofaktoren und Enzymen.
- 25 17. Verfahren zur Herstellung eines Polypeptides, umfassend das Züchten der Wirtszelle nach Anspruch 12 in einem geeigneten Kulturmedium, um so das Polypeptid zu produzieren.
- 30 18. Isoliertes MCP-Polypeptid aus *Corynebacterium glutamicum* oder ein Abschnitt davon.
19. Polypeptid nach Anspruch 18, wobei das Polypeptid an der Produktion einer Feinchemikalie beteiligt ist.
- 35 20. Isoliertes Polypeptid, umfassend eine Aminosäuresequenz, ausgewählt aus der Gruppe, bestehend aus den in Anhang B angegebenen Sequenzen.
- 40 21. Isoliertes Polypeptid, umfassend eine natürlich vorkommende allelische Variante eines Polypeptides, umfassend eine Aminosäuresequenz, ausgewählt aus der Gruppe, bestehend aus den in Anhang B angegebenen Sequenzen, oder einen Abschnitt davon.
- 45 22. Isoliertes Polypeptid nach einem der Ansprüche 18-21, das zudem heterologe Aminosäuresequenzen umfaßt.

23. Isoliertes Polypeptid, das von einem Nukleinsäuremolekül codiert wird, umfassend eine Nukleotidsequenz, die mindestens zu 50% homolog zu einer Nukleinsäure ist, ausgewählt aus der Gruppe, bestehend aus den in Anhang A angegebenen Sequenzen.
- 5 24. Isoliertes Polypeptid, umfassend eine Aminosäuresequenz, die mindestens zu 50% homolog zu einer Aminosäuresequenz ist, ausgewählt aus der Gruppe, bestehend aus den in Anhang B angegebenen Sequenzen.
- 10 25. Verfahren zur Herstellung einer Feinchemikalie, umfassend das Züchten einer Zelle, die einen Vektor nach Anspruch 12 enthält, so daß die Feinchemikalie produziert wird.
- 15 26. Verfahren nach Anspruch 25, wobei das Verfahren zudem den Schritt Gewinnen der Feinchemikalie aus der Kultur umfaßt.
27. Verfahren nach Anspruch 25, wobei das Verfahren zudem den Schritt Transfizieren der Zelle mit dem Vektor nach Anspruch 11 umfaßt, so daß eine Zelle erhalten wird, die den Vektor enthält.
- 20 28. Verfahren nach Anspruch 25, wobei die Zelle zur Gattung *Corynebacterium* oder *Brevibacterium* gehört.
- 25 29. Verfahren nach Anspruch 25, wobei die Zelle ausgewählt ist aus der Gruppe, bestehend aus: *Corynebacterium glutamicum*, *Corynebacterium herculis*, *Corynebacterium lilium*, *Corynebacterium acetoacidophilum*, *Corynebacterium acetoglutamicum*, *Corynebacterium acetophilum*, *Corynebacterium ammoniagenes*, *Corynebacterium fujiokense*, *Corynebacterium nitrilophilus*, *Brevibacterium ammoniagenes*, *Brevibacterium butanicum*, *Brevibacterium divaricatum*, *Brevibacterium flavum*, *Brevibacterium healii*, *Brevibacterium ketoglutamicum*, *Brevibacterium ketosoreductum*, *Brevibacterium lactofermentum*, *Brevibacterium linens*, *Brevibacterium paraffinolyticum* und den in Tabelle 3 angegebenen Stämmen.
- 30 30. Verfahren nach Anspruch 25, wobei die Expression des Nukleinsäuremoleküls von dem Vektor die Modulation der Produktion der Feinchemikalie bewirkt.
- 40 31. Verfahren nach Anspruch 25, wobei die Feinchemikalie ausgewählt ist aus der Gruppe, bestehend aus: organischen Säuren, proteinogenen und nichtproteinogenen Aminosäuren, Purin- und Pyrimidinbasen, Nukleosiden, Nukleotiden, Lipiden, gesättig-
- 45

ten und ungesättigten Fettsäuren, Diolen, Kohlehydraten, aromatischen Verbindungen, Vitaminen, Cofaktoren und Enzymen.

32. Verfahren nach Anspruch 25, wobei die Feinchemikalie eine Aminosäure ist.
33. Verfahren nach Anspruch 32, wobei die Aminosäure aus der Gruppe stammt, bestehend aus: Lysin, Glutamat, Glutamin, Alanin, Aspartat, Glycin, Serin, Threonin, Methionin, Cystein, Valin, Leucin, Isoleucin, Arginin, Prolin, Histidin, Tyrosin, Phenylalanin und Tryptophan.
34. Verfahren zur Herstellung einer Feinchemikalie, umfassend das Züchten einer Zelle, deren genomische DNA durch Einschluß eines Nukleinsäuremoleküls nach einem der Ansprüche 1-9 verändert worden ist.

CODIERENDE GENE NEUER PROTEINE VON *CORYNEBACTERIUM GLUTAMICUM*

Zusammenfassung der Offenbarung

5

Isolierte Nukleinsäuremoleküle, die als MCP-Nukleinsäuremoleküle bezeichnet werden und neue MCP-Proteine aus *Corynebacterium glutamicum* codieren, werden beschrieben. Die Erfindung stellt zudem Antisense-Nukleinsäuremoleküle, rekombinante Expressionsvektoren,

10

die MCP-Nukleinsäuremoleküle enthalten, und Wirtszellen, in die die Expressionsvektoren eingebracht worden sind, bereit. Sie stellt weiterhin isolierte MCP-Proteine, mutierte MCP-Proteine, Fusionsproteine, antigene Peptide und Verfahren zur Verbesserung der Produktion einer gewünschten Verbindung aus *C. glutamicum* be-

15

reit, die auf der genetischen Manipulation von MCP-Genen in diesem Organismus beruhen.

20

25

30

35

40

45